International Journal of Instruction *July 2026 ● Vol.19, No.3*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 215-238

Article submission code: 20250918141204
Received: 18/09/2025
Revision: 05/01/2026
Accepted: 14/01/2026
OnlineFirst: 03/04/2026

Psychometric Validity of an Instrument for Measuring Gender Gaps for Women in STEM Fields

Monserrat Vargas Jiménez
Universidad Autónoma de Yucatán, México, *monserrat.vargas@correo.uady.mx*

José Gabriel Domínguez Castillo
Universidad Autónoma de Yucatán, México, *jg.dominguez@correo.uady.mx*

Jose Manuel Ortiz-Marcos
University Granada, Spain, *jm.ortiz.marcos@ugr.es*

Gender equality and quality education are central pillars of international policies. However, women continue to face gaps in access, participation, and retention in STEM. This study aimed to validate the EMA-STEM scale to measure factors associated with these gaps among female university students. The validation was conducted in two phases: an exploratory factor analysis with 373 participants and a confirmatory factor analysis with 410 participants, all enrolled in STEM programs. The instrument, comprising 46 items and four dimensions—family support, gender bias, gender stereotypes, and gender inequity—was developed based on specialized literature and reviewed by a panel of experts. Statistical analyses revealed high reliability levels (α and ω > 0.80) and a robust factorial structure. The CFA model supported theoretically coherent associations among dimensions, highlighting family support as the most consistent. Finally, the EMA-STEM emerges as a valid and reliable tool to guide programs and policies aimed at reducing gender inequalities in STEM.

Keywords: family support, gender stereotypes, gender inequality, STEM, content validity

INTRODUCTION

Gender gaps in Science, Technology, Engineering, and Mathematics (STEM) persist worldwide. Recent reports from the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2024a) and UN Women (2025) highlight the inequalities women face not only in enrollment and graduation rates, but also in access to employment opportunities, leadership positions, and research and innovation funding. Additionally, women in STEM are more likely to experience gender-based violence and sexism compared to those in other areas of knowledge.

As a result, there has been a growing interest in understanding the conditions that influence women's choice, participation, and retention in STEM careers. This interest

Citation: Vargas-Jiménez, M., Domínguez-Castillo, J. G. & Ortiz-Marcos, J. M, (2026). Psychometric validity of an instrument for measuring gender gaps for women in STEM Fields. *International Journal of Instruction*, *19*(3), 215-238.

takes place within a global context where gender equality has become a strategic priority in educational, scientific, and development agendas (Economic Commission for Latin America and the Caribbean [ECLAC] and UN Women, 2024; UNESCO, 2024b; World Bank, 2024; World Economic Forum [WEF], 2025). Specifically, the 2030 Agenda recognizes education and gender equality as fundamental pillars for building more just societies (United Nations, 2023).

This reality—both in academic and social environments—does not stem from a single cause, but rather from a set of factors that women face when choosing male-dominated careers. Among the most significant elements are numerical underrepresentation, lack of support networks, unwelcoming academic environments, limited early promotion in STEM, and the reinforcement of gender stereotypes (Casad et al., 2021; Esteban-Ramiro et al., 2024). These stereotypes represent a set of beliefs shared by a society about the characteristics that men and women have or should have (Stewart et al., 2021), and they become more entrenched with age, directly influencing women's career choices (Coronado, 2022). This stereotypical dimension not only conditions academic culture but also extends into the broader sociocultural context.

In contexts of greater vulnerability, family dynamics play an even more decisive role in women's educational trajectories (Neri et al., 2022; Slattery et al., 2023). Domínguez-Castillo et al. (2024) note that university students face resistance within their own households, stemming from traditional structures that prioritize domestic and caregiving roles over professional training. These findings reaffirm that family support influences not only the emotional realm but also the material conditions that enable—or restrict—women's ability to remain in the education system (Lucena-Rodríguez et al., 2025).

Alongside the role of the family, educational experiences are also shaped by other dynamics that reproduce gender inequalities. Gender bias—unlike stereotypes—is reflected in practices that establish differentiated expectations, influence academic orientation, limit learning opportunities, and affect women's confidence in their abilities (Marchionni et al., 2018, p. 29). These biases are often linked to expectations around academic performance (Gallego & Casadiego, 2023), the way roles and activities are assigned in the classroom (Costello et al., 2024), or the promotion of interest in STEM (Dost, 2024).

In higher education, these biases are most evident in areas such as scholarship allocation, participation in research projects, or academic visibility, which forces female students to meet additional demands to validate their sense of belonging and legitimacy as women in these disciplines (Kinyota, 2021). Moreover, these biases mostly operate unconsciously and are reproduced through everyday practices, making it crucial to identify and become aware of existing barriers and biases in order to promote equity (Bordieri et al., 2024).

Beyond the reproduction of stereotypes and biases in family and school settings, gender gaps have structural implications. Gender inequality is reflected in the disparity between men and women in terms of opportunities and rights across various social spheres. In STEM, this exclusion is evidenced by the lack of recognition and support from authority figures, persistent wage gaps, the perception of lower competence attributed to women, reduced employment opportunities, and limited professional development (Neri et al., 2022). Furthermore, women's low participation in scientific production in STEM contributes to the invisibilization of their academic work, negatively impacting the advancement of scientific knowledge (Astegiano et al., 2019).

González-Cid (2024) conducted a study to document gender inequities in STEM. The study revealed the coexistence of a gender regime and a meritocratic academic regime which, operating together, produce, reproduce, and legitimize inequalities. It also found a tendency to overlook these gaps—even among female academics themselves—which encourages individual coping strategies rather than collective responses, thereby deepening the reproduction of inequalities in the educational system (Shahin et al., 2021). Thus, gender inequality is not limited to numerical representation but is embedded in normative and organizational structures that restrict women's participation, recognition, retention, and advancement in these disciplines (Lim et al., 2021).

Based on the above, academic literature has helped identify patterns, significant relationships between constructs, and explanatory models regarding the barriers women face in STEM. Instruments have also been developed to measure gender gap factors in higher education.

The EMA-STEM instrument is grounded in a multidimensional conceptual framework that understands gender gaps in STEM not as isolated individual phenomena, but as the result of interconnected social, cultural, and structural factors operating across women's educational trajectories. In this sense, the technical indicators of the instrument represent theoretically defined domains that capture key mechanisms through which gender inequality is produced and reproduced in STEM contexts (Dönmez, 2023).

The first indicator, family support, is conceptualized as a foundational socialization factor that shapes educational aspirations, self-efficacy, and career decision-making from early stages. Family beliefs, expectations, and the distribution of domestic responsibilities can either reinforce traditional gender roles or function as protective factors that encourage women's persistence in male-dominated fields (Piva & Rovelli, 2022). Prior research has shown that family environments play a decisive role in legitimizing—or discouraging—women's participation in STEM, particularly in contexts marked by cultural or socioeconomic vulnerability.

The second indicator, gender bias, refers to the presence of explicit or implicit practices that generate differentiated expectations and opportunities for women within academic and social environments. Unlike stereotypes, gender bias operates through everyday interactions, institutional norms, and evaluative processes that affect women's confidence, academic recognition, and access to resources. This indicator captures experiential dimensions of discrimination that directly impact women's sense of belonging and academic performance in STEM programs (Gregor et al., 2021).

The gender stereotypes indicator addresses the internalization of socially shared beliefs that associate scientific and technological competence with masculinity (Keller et al., 2022). These stereotypes are transmitted through media, educational narratives, and cultural representations, shaping perceptions of what careers are considered appropriate for women. The persistence of such stereotypes has been widely documented as a barrier that limits interest, self-identification, and long-term engagement of women in STEM fields.

Finally, gender inequity is conceptualized as a structural dimension that reflects systemic inequalities in access to opportunities, recognition, and professional advancement (Yokoyama et al., 2024). This indicator goes beyond individual perceptions to capture broader institutional and labor-market dynamics, such as unequal valuation of women's work, restricted career progression, and tolerance of discriminatory practices. Gender inequity thus represents the macro-level context in which family dynamics, biases, and stereotypes converge.

Together, these four indicators provide a coherent and theoretically grounded framework that guides the construction of the EMA-STEM instrument. Their integration allows for a comprehensive assessment of women's experiences in STEM, linking subjective perceptions with structural conditions and clarifying the analytical direction of the study.

Although these works offer valuable frameworks for understanding the phenomenon, they often omit situated aspects—i.e., they are aimed at mixed populations or do not focus exclusively on STEM—which hinders the ability to capture the specificities of women's trajectories in these fields (Table 1).

Table 1
Empirical and theoretical references for the construction of the EMA-STEM scale

Author(s)	Country	Population	Domains	Validity and reliability
Yang & Gao (2021)	China	University students	1. Achievement Motivation 2. Parents', Professional, Expectations 3. University Professors', Gender Stereotypes 4. Cultural Gender Stereotypes 5. Attitudes, Toward, Gender, Roles	Domain 1 Cronbach's Alpha (α) = .806 Domain 3 α = .568 Domain 4 α = .751 Domain 5 α = .867
Beroíza-Valenzuela et al. (2024)	Chile	University students	1. Coping in University Students 2. Academic Beliefs 3. Family Academic Support 4. University Socialization 5. Collaborative Learning Experience 6. Gender Role Beliefs	Literature review Expert interviews Delphi technique with nine specialists in gender and psychometric validation Exploratory and confirmatory factor analysis α per dimension: 1 = 0.72; 2 = 0.72; 3 = 0.62; 4 = 0.68; 5 = 0.72; 6 = 0.78 McDonald's Omega (ω): between 0.69 and 0.80 per dimension Delphi Agreement (Cohen's kappa): 0.85
Pownall & Heflick (2024)	United Kingdom	University students	1. Self-Perceived Gender Typicality 2. Approval of Gender Stereotypes 3. Satisfaction with Gender Role 4. Sexual Role Inventory 5. Academic Achievements	Domain 1 α (female version) = .83 α (male version) = .80 Domain 4 α (female version) = .83 α (male version) = .88
González-Rogado et al. (2025)	Spain	University students	1. Identification 2. Context 3. Perception-Gender Equality-Gender Gap-Gender Stereotypes	Validated Instrument GENCE 2.0 (Gender Perspective in Computer Engineering 2.0) α
Verdugo-Castro et al. (2025)	Spain	University students	1. Interest 2. Perception and Self-Perception 3. Gender Ideology 4. Attitudes 5. Expectations about Science	α per dimension: 1 = 0.741; 2 = 0.746; 3 = 0.726; 4 = 0.645; 5 = 0.761

Note. Own elaboration.

In this context, it is relevant to develop psychometrically validated instruments specifically designed to measure the factors influencing women's experiences —not in comparison to men— regarding their participation in STEM. The aim of this study is to validate an instrument that identifies the main factors associated with women's participation and retention in STEM fields. The Survey for Women in STEM Fields (EMA-STEM) is targeted at women enrolled in these disciplines, allowing for a deeper exploration of the issue and addressing the limitations in design, scale, and focus found in previous studies. It includes 46 items across four domains —family support, gender bias, gender stereotypes, and gender inequity— developed from specialized literature and statistical analysis, aligned with the structural and subjective conditions that characterize women's experiences in STEM.

METHOD

Methodology

The EMA-STEM consists of two sections. The first corresponds to sociodemographic data; the second, to contextual factors in STEM. Table 2 presents the structure of the instrument according to its sections, specifying the evaluated domains, item distribution, and response scale.

Table 2
EMA-STEM Instrument Design

Section	Domain	Number of Items	Items	Scale
Personal Data	N/A	N/A	N/A	N/A
Contextual Factors in STEM	Family Support	17	A1 – A17	1 Strongly Disagree 2 Disagree 3 Neither Agree nor Disagree 4 Agree 5 Strongly Agree
	Gender Bias	10	B1 – B10	
	Gender Stereotypes	8	C1 – C8	
	Gender Inequity	11	D1 – D11	

Note. (Bueno & Duchemin, 2022; Block & Schmader, 2025; O'Connor, 2023; Shanks et al., 2024).

The "family support" domain consists of items that complete the phrase "My parents…" and includes statements about generalized beliefs within the family environment, gender-based task assignments, perceptions of careers deemed appropriate for men or women, and the distribution of family responsibilities (Bueno & Duchemin, 2022). The "gender bias" domain includes items that complete the phrase "As a student…" and address experiences of discrimination that women may face as students in family, social, academic, or professional settings (Shanks et al., 2024). The "gender stereotypes" domain features items that complete the phrase "As a woman…" and includes statements about societal expectations, femininity in advertising, and the influence of traditional narratives on professional aspirations (Block & Schmader, 2025). Lastly, the "gender inequity" domain comprises items that assess disparities and barriers to accessing professional opportunities (O'Connor, 2023).

Instrument Design

The design of the EMA-STEM instrument was developed in three phases. First, a comprehensive review of recent literature on gender gaps in STEM within higher education was conducted, providing the theoretical and empirical foundations for the scale's construction. In the second stage, content validity was established through expert judgment: a panel of three specialists in gender studies and one expert in STEM fields evaluated each item in terms of clarity, coherence, and relevance, reaching consensus when items received favorable ratings across all three attributes. Finally, the instrument underwent psychometric evaluation, which included an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA). The EFA allows for the identification of the underlying structure of the items and their grouping into coherent factors according to theoretical expectations, while the CFA is used to confirm whether the empirical data adequately fit the proposed factorial model, ensuring the structural validity of the scale.

Phase 1: Literature Review

The literature review drew upon recent empirical studies on gender gaps in STEM in higher education across different contexts. Studies were selected based on theoretical and methodological validity, as well as their ability to measure the variables of interest (Table 1).

Phase 2: Item Drafting

The initial drafting of items was based on the findings from Phase 1, prioritizing clarity and brevity in the wording, as well as conceptual coherence and relevance of the included domains. An accessible language was used to ensure participants could fully understand the items and provide clear, unambiguous responses.

Phase 3: Content Validity

Content validity was assessed through expert judgment (Maxim, 2002). Four expert judges were selected based on specific criteria: three gender studies specialists contributed to the theoretical and conceptual evaluation of the instrument from a specialized perspective, and one expert in STEM fields provided input on the relevance of items in relation to the disciplinary context. The items were rated on a four-point scale according to three attributes: clarity, coherence, and relevance. An item was considered to have consensus when it received a rating of three or four in all three attributes from the judges. These items formed the final version of the instrument (Table 3).

Table 3
Results of the expert judgment regarding the Content Validity Coefficient (CVC)

ÍTEMS	25	50	75	P 75-25	Decision
1	4,00	4,00	4,00	0,00	Accept
2	4,00	4,00	4,00	0,00	Accept
3	4,00	4,00	4,00	0,00	Accept
4	4,00	4,00	4,00	0,00	Accept
5	4,00	4,00	4,00	0,00	Accept
6	2,28	4,00	4,00	1,72	Modify
7	2,78	4,00	4,00	1,22	Modify
8	2,25	3,50	4,00	1,75	Modify
9	2,25	4,00	4,00	1,75	Modify
10	4,00	4,00	4,00	0,00	Accept
11	4,00	4,00	4,00	0,00	Accept
12	4,00	4,00	4,00	0,00	Accept
13	4,00	4,00	4,00	0,00	Accept
14	4,00	4,00	4,00	0,00	Accept
15	4,00	4,00	4,00	0,00	Accept
16	4,00	4,00	4,00	0,00	Accept
17	4,00	4,00	4,00	0,00	Accept
18	4,00	4,00	4,00	0,00	Accept
19	4,00	4,00	4,00	0,00	Accept
20	4,00	4,00	4,00	0,00	Accept
21	4,00	4,00	4,00	0,00	Accept
22	4,00	4,00	4,00	0,00	Accept
23	4,00	4,00	4,00	0,00	Accept
24	4,00	4,00	4,00	0,00	Accept
25	4,00	4,00	4,00	0,00	Accept
26	4,00	4,00	4,00	0,00	Accept
27	4,00	4,00	4,00	0,00	Accept
28	4,00	4,00	4,00	0,00	Accept
29	4,00	4,00	4,00	0,00	Accept
30	4,00	4,00	4,00	0,00	Accept
31	4,00	4,00	4,00	0,00	Accept
32	2,25	4,00	4,00	1,75	Modify
33	3,00	4,00	4,00	1,00	Modify
34	2,25	3,50	4,00	1,75	Modify
35	4,00	4,00	4,00	0,00	Accept
36	4,00	4,00	4,00	0,00	Accept
37	4,00	4,00	4,00	0,00	Accept
38	4,00	4,00	4,00	0,00	Accept
39	4,00	4,00	4,00	0,00	Accept
40	4,00	4,00	4,00	0,00	Accept
41	4,00	4,00	4,00	0,00	Accept
42	4,00	4,00	4,00	0,00	Accept
43	4,00	4,00	4,00	0,00	Accept
44	4,00	4,00	4,00	0,00	Accept
45	4,00	4,00	4,00	0,00	Accept
46	4,00	4,00	4,00	0,00	Accept

Instrument Validation

Instrument validation was conducted in two complementary phases. In Phase 1, an Exploratory Factor Analysis (EFA) was performed to identify the latent structure underlying the instrument's items. In Phase 2, a Confirmatory Factor Analysis (CFA) was carried out to test the fit of the theoretical model derived from the EFA. Additionally, Cronbach's alpha (α) and McDonald's omega (ω) reliability coefficients were calculated to assess internal consistency and determine whether the instrument maintains a coherent and stable structure in which all items contribute equally to the measurement of each domain (Colorado et al., 2024).

Finally, construct validity—understood as the extent to which an instrument accurately measures the theoretical concept it aims to assess (Hair et al., 2019)—was examined through a combined EFA and CFA approach, allowing for both empirical exploration and confirmation of the instrument's factorial structure.

Population

This study employed a non-probabilistic purposive sampling method. In the first phase (EFA), a sample of 373 participants was analyzed (N = 373). Data collection took place in March 2025 using QuestionPro, and the data were processed in RStudio version 2025.05.0+496. The participants were female students enrolled in STEM academic programs, including the following disciplines: sustainable agriculture, agroecology, nursing, environmental engineering, biomedical engineering, civil engineering, food engineering, biotechnology engineering, industrial engineering, petroleum engineering, chemical engineering, mathematics, mechanics, medicine, nutrition, and information and communication technologies. Participants in this phase were aged 18 to 26 years (M = 20.08; SD = 1.818).

The second phase (CFA) included a sample of 410 female participants (N = 410). Data collection was conducted during May and November 2025, also via QuestionPro, and data were processed using Jamovi (version 26) and Rstudio (versión 2025.09.2+418). As in the previous phase, the sample consisted of female students enrolled in STEM academic programs in the same disciplines mentioned above. Participants in this phase ranged in age from 18 to 26 years (M = 20.20; SD = 1.861). The findings are sufficient and useful for the formulation of research recommendations.

With respect to the characteristics of the study subjects, it should be noted that the participants come from different regions of Mexico and belong predominantly to socially vulnerable groups and traditional family structures. These families are often characterized by limited economic resources, restricted access to educational opportunities, and strong cultural and familial ties. Such contextual factors play a significant role in shaping the participants' experiences and behaviors analyzed in this study. Therefore, the research findings should be interpreted within this broader social and familial context, as the family environment constitutes a key element influencing the observed outcomes.

The adequacy of the sample size in both phases was determined based on established statistical criteria and contextual population estimates. In Mexico, official reports from institutions such as the Ministry of Public Education (SEP), the Mexican Institute for Competitiveness (IMCO), the National Autonomous University of Mexico (UNAM), and Secretariat of Science, Humanities, Technology and Innovation (Secihti) indicate that women represent approximately 27–30% of total enrollment in STEM fields at the higher education level, with significant variation by region and discipline. These data, disseminated through official publications and specialized media reports, also show a progressive decrease in female participation at postgraduate and research levels.

Based on this population context, sample size estimation was conducted assuming a large population of female students in STEM fields. According to OECD (2022), women represented approximately 32% to 35% of enrollment and graduates in engineering, manufacturing, and construction (EMC) programs in Mexico, with a total of 494,753 women enrolled in STEM-related programs nationwide. Considering this population size, a confidence level of 99% and a maximum margin of error of 5% were applied. The sample size calculation indicated that a minimum of 355 participants was required for the study; accordingly, the sample sizes obtained across both phases exceeded this threshold, meeting commonly recommended criteria for exploratory and confirmatory factor analyses. Although the sampling strategy was non-probabilistic, the achieved sample sizes are considered statistically adequate to support the robustness and stability of the instrument's factorial structure.

The study received approval from the Academic Committee of the Master's Program in Educational Research (CAMIE), Faculty of Education, Autonomous University of Yucatán, Mexico, on May 29, 2025. Informed consent was obtained digitally, ensuring participants' understanding, voluntariness, and confidentiality.

FINDINGS

Exploratory Factor Analysis (EFA)

The EFA identified four distinct dimensions within the scale. The analysis used principal component extraction with Oblimin rotation, based on the expectation of correlations among factors (Mineiro et al., 2025). Sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. Only items with communalities equal to or greater than 0.30 were retained (Tomé-Fernández et al., 2019). Multiple criteria were used to determine the number of factors and select items for removal, establishing that items with factor loadings or communalities below 0.30 would be eliminated. After item screening based on communalities and factor loadings, the EFA was re-estimated to confirm the stability of the four-factor solution. The final structure remained theoretically coherent and statistically robust (Hair et al., 2019; Salahshouri & Mohamadian, 2024).

Table 4
Extraction of Communalities in Exploratory Factor Analysis (EFA)

Items	Extraction	Items	Extraction
1	,644	24	,636
2	,605	25	,787
3	,837	26	,766
4	,841	27	,832
5	,793	28	,756
6	,675	29	,808
7	,656	30	,632
8	,811	31	,776
9	,712	32	,696
10	,685	33	,640
11	,860	34	,679
12	,706	35	,669
13	,786	36	,785
14	,623	37	,650
15	,761	38	,882
16	,807	39	,858
17	,735	40	,762
18	,643	41	,880
19	,734	42	,630
20	,616	43	,646
21	,726	44	,900
22	,645	45	,846
23	,677	46	,607

This factor analysis was supported by an adequate Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy (KMO = 0.93) and a significant Bartlett's test of sphericity (χ^2 = 7. 225,108, p= .000), indicating that the correlation matrix was suitable for factor analysis. The four extracted factors accounted for 63,16% of the total variance. Based on the rotated component matrix, four domains were identified (Table 5).

Table 5
Rotated Component Matrix

Items	Factors 1	2	3	4	Dimension	Cronbach	Omega de McDonald (ω)
					Family Support	0.92	0.93
A1	,841						
A2	,788						
A3	,777						
A4	,776						
A5	,739						
A6	,721						
A7	,687						
A8	,671						
A9	,659						
A10	,611						
A11	,793						
A12	,756						
A13	,751						
A14	,744						
A15	,740						
A16	,731						
A17	,729						
					Gender Bias	0.87	0.90
B1		,870					
B2		,878					
B3		,862					
B4		,829					
B5		,685					
B6		,328					
B7		,409					
B8		,635					
B9		,624					
B10		,623					
					Gender Stereotypes	0.80	0.83
C1			,420				
C2			,754				
C3			,669				
C4			,663				
C5			,678				
C6			,830				
C7			,729				
C8			,727				
					Gender Inequity	0.85	0.94
D1				,889			
D2				,888			
D3				,686			
D4				,397			
D5				,639			
D6				,762			
D7				,823			
D8				,581			
D9				,662			
D10				,723			
D11				,581			
Eigenvalues	11,21	3,97	1,96	1,17			
% Explained variance	38,66	13,69	6,77	4,04			
% Cumulative explained variance	38,66	52,35	59,12	63,16			
KMO	0.93						
Bartlett's test of sphericity	X^2=7. 225,108 Sig. ≤ 0.00						

Notes. Extraction method: Principal Component Analysis. Rotation method: Oblimin rotation with Kaiser normalization.

Following the Exploratory Factor Analysis, the final version of the instrument consisted of 46 items distributed across four dimensions, as determined by the rotated component matrix.

Factor 1: Family Support

Assesses the role of the family in women's participation in STEM. It examines traditional family beliefs, gender expectations (e.g., preference for women to be housewives), and the resources or barriers present in the family environment. It comprises 17 items and has excellent internal consistency ($\alpha = 0.92$).

Factor 2: Gender Bias

Explores how biases in the school and social environments impact women's self-esteem and performance. It includes 10 items addressing imposed household duties, ten comments, classroom discrimination, and limited access to educational infrastructure. This dimension shows acceptable reliability ($\alpha = 0.87$).

Factor 3: Gender Stereotypes

Reflects how media, advertising, and popular culture influence the perception that STEM fields are "for men" and how this discourages or limits women from an early age. It includes 8 items with high internal consistency ($\alpha = 0.80$).

Factor 4: Gender Inequity

Examines structural inequalities—such as workplace discrimination, sexual harassment, lower recognition of women's work, and the notion that their role is confined to the home—that women face in STEM. This domain consists of 11 items and shows high reliability ($\alpha = 0.85$). Table 5 presents the results of the EFA and reliability analysis, showing that the EMA-STEM yields acceptable, good, or excellent levels of measurement consistency, indicating that the scale's assessments are coherent and reliable.

It is worth noting that, in addition to Cronbach's α, each dimension was also evaluated using McDonald's ω coefficient, since—as highlighted by authors such as Yamamoto (2023)—this measure, which is based on factor loadings, is currently considered a more precise and reliable alternative. The results obtained are consistent with previous studies (Guilarte et al., 2025; Yang & Chittoori, 2022) that address similar dimensions (Table 5).

Confirmatory Factor Analysis (CFA)

Following the EFA, a CFA was conducted to assess the adequacy of the indicators in representing the latent constructs (Felicetti & Cabrera, 2022). For this analysis, the Minimum Discrepancy statistic (CMIN) with a chi-square distribution was examined (see Table 6). Figure 1 presents a Structural Equation Model (SEM) composed of four interrelated latent factors, labeled factor 1 (F1), factor 2 (F2), factor 3 (F3), and factor 4 (F4). Each factor groups a set of observable items corresponding to participant responses on a 46-item questionnaire, identified as A1–A17, B1–B10, C1–C8, and D1–

D11. The assignment of items to factors was based on both the factor loading values obtained and the theoretical coherence of the instrument.

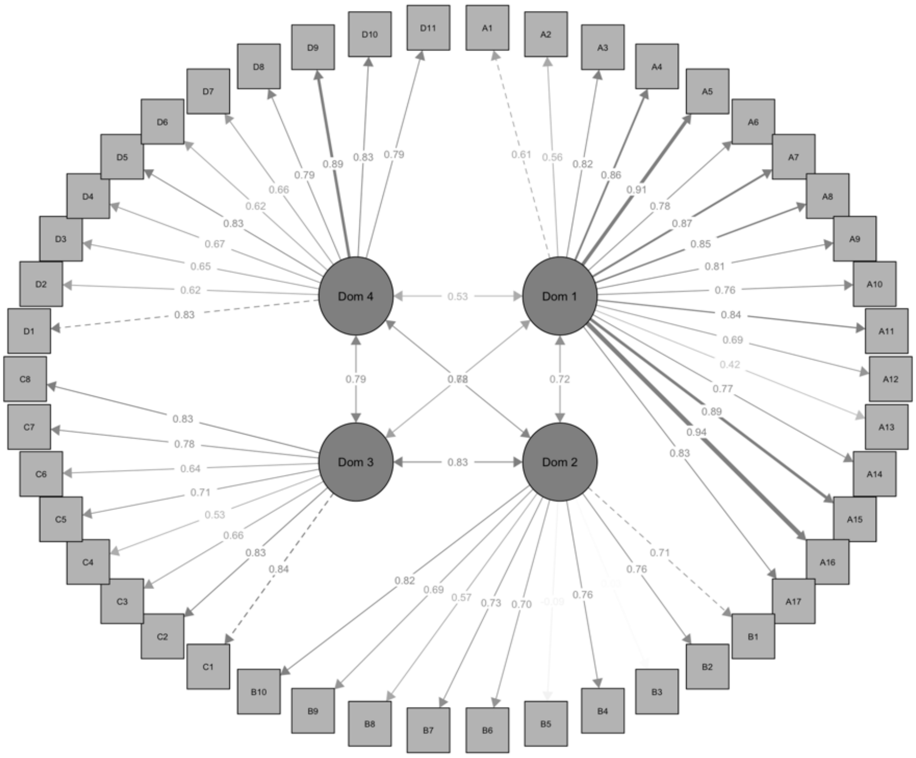

Figure 1
Structural Equation Model of the EMA-STEM Instrument

A structural equation model was estimated with the aim of evaluating the validity of the questionnaire and the associations between its latent dimensions. The model included four latent constructs (Dom1, Dom2, Dom3, and Dom4), measured respectively by item sets A, B, C, and D. The analysis was performed on standardized factor loadings and correlations between factors, interpreting the SEM model with correlated factors.

First, the measurement model showed a pattern of mostly high factor loadings, which supports the convergent validity of the constructs. In the case of Dom1, the standardized factor loadings ranged from approximately 0.42 to 0.94, with most items showing high saturations ($\lambda \geq 0.70$), indicating that these indicators adequately represent the latent construct. However, at least one item with a low loading ($\lambda \approx 0.42$) was identified, but it was retained due to its theoretical relevance, as it represents subtle manifestations of biases and micro-interactions widely documented in the literature on gender in STEM. Furthermore, its inclusion allows for comparability with previous applications of the instrument.

The Dom2 construct presented heterogeneous standardized factor loadings. While several indicators showed moderate-to-high loadings (e.g., B1–B2, B4, B6–B10), two items exhibited very low standardized loadings (B3: $\lambda = 0.028$, $p = .530$; B5: $\lambda = -0.085$, $p = .095$). This pattern may reflect item-specific characteristics, such as subtle or context-dependent expressions of gender bias, which have been reported in prior research as more difficult to capture through self-report measures. The Dom3 showed standardized factor loadings ranging from approximately 0.53 to 0.84, with most items reaching acceptable or high levels of saturation, although one of the indicators was at the lower recommended limit. Finally, Dom4 showed solid performance, with standardized factor loadings ranging from approximately 0.62 to 0.89, indicating adequate internal consistency and a correct definition of the latent construct.

Overall, the factor loadings suggest high composite reliability for the four domains (CR > 0.70). While AVE values exceeded the recommended threshold of 0.50 for three domains, the Gender Bias factor showed a slightly lower AVE, which was considered acceptable given its high composite reliability and strong theoretical grounding. These results indicate that the items explain a substantial proportion of the variance in their respective latent factors.

In relation to the structural model, the standardized correlations between the four domains were positive and statistically high, with values ranging from approximately 0.53 to 0.88. These associations indicate that the constructs are closely related, which is consistent with their belonging to the same conceptual framework.

In summary, the proposed SEM model reveals the existence of four interrelated dimensions that solidly and coherently structure the construct measured by the 46 items. The associations among latent factors reflect a multidimensional and interconnected configuration of the phenomenon, consistent with theoretical approaches that describe structured or systemic processes related to self-evaluation and perception. As a preliminary absolute fit indicator, the Minimum Discrepancy Chi-Square statistic (CMIN) was examined (Table 6).

Table 6
CMIN as a model fit indicator in CFA

	NPAR	CMIN	DF	P	CMIN/DF
Default model	236	4164.230	983	< 0.001	4.24
Saturated model	2277	0.000	0		
Independence model		125.511,883	1035	< 0.001	121.27

Note. CMIN = Minimum Discrepancy Chi-square Statistic.

As noted in previous studies (Beribisky & Hancock, 2023; Papagiannopoulou et al., 2023), Chi-Square–based tests can be unreliable. Consequently, alternative indices were incorporated—the Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA)—to assess the fit between observed and predicted covariance matrices (Paige et al., 2019).

The CFI, TLI, and RMSEA values obtained indicated an acceptable model fit (Fan et al., 2025). While values close to 1.00 in CFI and TLI suggest a good fit (Jin & Chen, 2024), the results in this study (CFI = 0.974; TLI = 0.973) reflect a moderate fit (Table 7).

Table 7
Goodness of fit EMA-STEM

χ^2 (gl)	p	CFI	TLI	RMSEA	RMSEA IC 90%	SRMR
4164.23 (983)	< .001	0.974	0.973	0.08	0.086–0.092	0.087

Note. TLI: = Tucker–Lewis index; CFI = comparative

In this context, RMSEA (Root-Mean-Square Error of Approximation) is recognized as one of the most informative indicators for evaluating SEM (Counsell et al., 2019; Morata-Ramírez et al., 2015). This index accounts for degrees of freedom when determining goodness of fit, making it sensitive to the number of estimated parameters in the model (Cortes et al., 2021). In other words, RMSEA values tend to decrease as degrees of freedom or sample size increase (Beribisky & Hancock, 2023). RMSEA values below .05 are considered indicative of good fit, whereas values ranging from 0.05 to 0.08 indicate reasonable fit (Browne & Cudeck, 1993); values slightly above this threshold may be observed in complex models with multiple factors and ordinal indicators.

RMSEA values suggested a more conservative assessment of model fit (RMSEA = 0.08, 90% CI [0.086–0.092]). This pattern of results indicates a mixed fit, characterized by strong incremental fit indices alongside a more conservative absolute fit index. Such discrepancies have been documented in large, multidimensional measurement models with ordinal indicators, where RMSEA is known to be sensitive to model complexity and degrees of freedom.

Taken together, the fit indices suggest that the proposed CFA model provides an acceptable representation of the data, while acknowledging areas of misfit commonly observed in complex social science measurement models.

Instrument Reliability

Once the CFA was completed and the structure of the instrument confirmed, internal consistency tests (α and ω) were performed again, yielding excellent reliability for each of the dimensions that make up the scale (Table 8).

Table 8
Reliability and validity of the EMA-STEM instrument (CFA)

Domain	α	α ordinal	ω	AVE
Family support	0.93	0.96	0.94	0.62
Gender biases	0.75	0.82	0.86	0.42
Gender stereotypes	0.85	0.90	0.97	0.54
Gender inequality	0.88	0.91	0.92	0.56

Note. Own elaboration.

Although AVE values exceeded the recommended threshold of 0.50 for three of the four factors, the Gender Bias dimension showed a slightly lower AVE (0.42). However, this value was considered acceptable given the high composite reliability of the factor (ω = 0.86) and its strong theoretical grounding. According to Hair et al. (2019), AVE values below 0.50 may be tolerated in established constructs when composite reliability is adequate and the construct is conceptually well-defined, supporting the convergent validity of this dimension.

Test–Retest Reliability

Test–retest reliability was evaluated using responses collected at two points in the study: the EFA and the CFA. Since the samples in each phase were different, a paired subset of cases was selected to compare responses across both moments. The results indicated excellent temporal stability of the instrument. The Intraclass Correlation Coefficient (ICC) was 0.93 (95% CI = 0.86 – 0.95), confirming excellent consistency across both applications. These results support the temporal reliability of the questionnaire and reinforce its usefulness as a consistent tool to evaluate factors associated with gender gaps in STEM careers. Finally, the model demonstrated an acceptable level of congruence and supports the hypothesis of the construct's multidimensionality. The EMA-STEM was established with 46 items distributed across four latent dimensions.

DISCUSSION

The results confirm the validity and reliability of the EMA-STEM instrument for measuring the factors that influence women's participation and retention in STEM careers. This finding aligns with the need for robust, context-sensitive psychometric instruments, particularly when measuring variables related to gender and structural barriers (Papagiannopoulou et al., 2023; Tomé-Fernández et al., 2019). It is also highly relevant in Latin American contexts, where deep gender inequalities persist in higher education and in scientific and technological fields (ECLAC & UN Women, 2024; Vargas Jiménez et al., 2025). The Global Gender Gap Report 2025 highlights that although significant progress has been made over the past two decades in narrowing the gender gap, women still score low in economic participation and opportunities, including access to and advancement in careers such as STEM (WEF, 2025).

Validity is framed within an emerging line of research that seeks to incorporate intersectional and context-based perspectives into the design of educational scales. Tomé-Fernández et al. (2019), in their study on cyberbullying in religiously and culturally diverse contexts, demonstrate that identity variables shape school experiences, especially among vulnerable populations. In this same logic, the EMA-STEM acknowledges that gender does not act in isolation but rather intersects with other social factors.

Chacón-Cuberos et al. (2025), Papagiannopoulou et al. (2023), and Lendínez-Turón et al. (2023) emphasize that a psychometric instrument cannot be disconnected from the social realities it aims to measure. Along these lines, Domínguez-Castillo et al. (2025) stress that unequal access to technological resources, teacher attitudes, and sociocultural conditions significantly influence women's educational experiences—determinants that must be considered when designing context-sensitive instruments. In this regard, EMA-STEM not only measures individual perceptions but also reflects broader systems of exclusion, making it a theoretical and methodological innovation in gender and STEM studies.

The factorial structure, supported by both EFA and CFA, revealed four domains—family support, gender bias, gender stereotypes, and gender inequity—consistent with the specialized literature (Casad et al., 2021; Lim et al., 2021; Marchionni et al., 2018; Neri et al., 2022; Slattery et al., 2023; Verdugo-Castro et al., 2025). This multidimensional approach allows for capturing the specific trajectories and experiences of women in STEM, as also evidenced in recent studies, such as the TRi-STEM Scale, which integrates affective, cognitive, and attitudinal dimensions (Papagiannopoulou et al., 2023). Moreover, EMA-STEM incorporates factors not considered in other scales and achieves outstanding levels of internal consistency was good to excellent across dimensions ($\omega = 0.80–0.94$), comparable to instruments such as KABIP-S (Lendínez-Turón et al., 2023), designed to assess knowledge, attitudes, behaviors, and participation intentions regarding the Sustainable Development Goals.

A detailed analysis of the dimensions showed that family support obtained the highest internal reliability ($\alpha = 0.96$; $\omega = 0.94$), reaffirming the decisive role of the family in guiding women's vocational choices toward traditionally male-dominated fields (Neri et al., 2022; Slattery et al., 2023). The high consistency of the gender bias domain highlights the influence of normalized practices that affect women's sense of belonging in STEM (Marchionni et al., 2018; Kinyota, 2021). Similarly, gender stereotypes align with studies showing how sociocultural imaginaries shape professional aspirations from early ages (Coronado, 2022; Esteban-Ramiro et al., 2024). Finally, the gender inequity domain revealed the persistence of structural barriers as mechanisms of systemic exclusion in women's professional trajectories (Astegiano et al., 2019; González-Cid, 2024). Although the confirmatory model indices reflected a moderate fit (CFI = 0.974; TLI = 0.973), these values are acceptable within factorial analysis standards, particularly given the sample size. Nevertheless, caution is recommended when interpreting the results, along with replications using larger and more diverse samples that would allow for model generalization, as well as complementary analyses considering longitudinal trajectories, cultural biases, and other intersectional factors. Such future research could further refine the scale's sensitivity to capture specific realities.

CONCLUSIONS

The aim of this study was to examine and validate the psychometric properties of the EMA-STEM scale through exploratory and confirmatory factor analyses. The exploratory factor analysis was re-estimated using a larger sample, resulting in a more stable and theoretically coherent factor solution. The final structure comprised four domains, jointly explaining 63.16% of the total variance, a value considered acceptable for instruments assessing complex psychoeducational constructs.

Based on the revised exploratory results, a partial reduction of items was implemented, particularly in those dimensions initially containing more than ten indicators. This refinement improved the balance across factors and reduced model saturation while preserving the conceptual integrity of the instrument. The retained items exhibited standardized factor loadings predominantly above 0.60, with many exceeding 0.70, indicating an adequate representation of the underlying latent constructs.

The confirmatory factor analysis of the adjusted model demonstrated a satisfactory overall fit, with goodness-of-fit indices meeting commonly accepted thresholds (CFI and TLI $\geq$ 0.90; RMSEA and SRMR $\leq$ 0.08). The majority of standardized factor loadings were statistically significant, supporting the adequacy of the revised measurement model and resolving the inconsistencies previously observed between the exploratory and confirmatory analyses.

Regarding reliability, the EMA-STEM scale showed good to excellent internal consistency across the four dimensions. Both Cronbach's alpha and McDonald's omega coefficients exceeded 0.80, reaching values close to or above 0.90 in some domains. In addition, indicators of convergent validity were satisfactory, with average variance extracted (AVE) values were above 0.50 for three domains, while the Gender Bias factor showed a lower AVE (0.42) that was considered acceptable given adequate composite reliability.

Although high correlations were observed among some factors ($r > 0.70$), suggesting a strong conceptual relationship between the dimensions, these findings are consistent with the theoretical framework underlying the EMA-STEM construct. Nevertheless, further research is warranted to continue examining discriminant validity and to explore alternative modeling approaches, such as second-order factor structures.

Overall, the numerical evidence obtained after the revisions supports the factorial validity, reliability, and internal consistency of the EMA-STEM scale. The instrument thus represents a psychometrically sound tool for assessing the proposed dimensions in STEM educational contexts. Future studies should extend this work by examining external and predictive validity, including relationships with variables such as STEM self-efficacy, career interest, academic persistence, and academic performance, in order to further enhance the practical relevance of the scale.

REFERENCES

Astegiano, J., Sebastianorte-González, E., & Castanho, C. (2019). Unravelling the gender productivity gap in science: a meta-analytical review. *Royal Society Open Science, 6*, 1–12. http://doi.org/10.1098/rsos.181566

Beribisky, N., & Hancock, G. R. (2023). Comparing RMSEA-Based Indices for Assessing Measurement Invariance in Confirmatory Factor Models. *Educational and Psychological Measurement, 84*(4), 716–735. https://doi.org/10.1177/00131644231202949

Beroíza-Valenzuela, F., Salas-Guzmán, N., & Huepe, D. (2024). Gender disparities in higher education: development and validation of the FACT-GÉN instrument. *Frontiers Education, 9*. https://doi.org/10.3389/feduc.2024.1456085

Block, K., & Schmader, T. (2025). Me, myself, and my stereotypes: does retraining gender stereotypes change men's self-concept? *Self and Identity, 24*(4), 326–358. https://doi.org/10.1080/15298868.2025.2477003

Bordieri, M. J., Waddill, P. J., Zhang, Q., McCarthy, M. L., Fuller, C., & Balthrop, D. (2024). Exploring the stability of the gender gap in faculty perceptions of gender climate at a rural regional university. *PLoS ONE, 19*(4). https://doi.org/10.1371/journal.pone.0301285

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. En K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Bueno, P., & Duchemin, M. H. (2022). Contribution of Psychological Entrepreneurial Support to the Strengthening of Female Entrepreneurial Intention in a Women-Only Incubator. *M@n@gement, 25*(4), 64–79. https://doi.org/10.37725/mgmt.v25.4556

Casad, B. J., Franks, J. E., Garasky, C. E., Kittleman, M. M., Roesler, A.C., Hall, D. Y., & Petzel, Z. W. (2021). Gender inequality in academia: Problems and solutions for women faculty in STEM. *Journal of Neuroscience Research, 99*, 13–23. https://doi.org/10.1002/jnr.24631

Chacón-Cuberos, R., Rodríguez-Sabiote, C., Expósito-López, J., Olmedo-Moreno, E., Serrano-García, J., & Hortas-Aliaga, O. (2025). Development of an employability thinking scale for use with young people in training: Exploratory and confirmatory factor analysis. *Systems, 13*(6), 479. https://doi.org/10.3390/systems13060479

Colorado, J., Romero, M., Salazar, M., Cabrera, G., y Castillo, V. (2024). Análisis Comparativo de los Coeficientes Alfa de Cronbach, Omega de McDonald y Alfa Ordinal en la Validación de Cuestionarios. *Estudios y Perspectivas: Revista Científica y Académica, 4*(4), 2738–2755. https://doi.org/10.61384/r.c.a.v4i4.836

Comisión Económica para América Latina y el Caribe y ONU Mujeres. (2024). *La Agenda 2030 para el Desarrollo Sostenible y la Agenda Regional de Género en América Latina y el Caribe: indicadores de género a 2023*. Santiago de Chile: CEPAL. https://repositorio.cepal.org/entities/publication/d884224f-bd0a-4758-9ab8-98dc791abbd8

Coronado, C. (2022). "¿Por qué tan pocas?": un proyecto audiovisual para visibilizar el trabajo de las científicas y tecnólogas españolas. *Investigaciones feministas, 13*(2), 613–623. https://dx.doi.org/10.5209/infe.80307

Cortes, S., Pineda, H., & Geverola, I. J. (2021). A confirmatory factor analysis of teacher's competence in action research (TCAR) questionnaire. *Advanced Education, 8*(19), 103–113. https://doi.org/10.20535/2410-8286.241148.

Costello, R., Salehi, S., Ballen, C., & Burkholder, E. (2023). Pathways of opportunity in STEM: comparative investigation of degree attainment across different demographic groups at a large research institution. *International Journal of STEM Education, 10*(46). https://doi.org/10.1186/s40594-023-00436-5

Counsell, A., Cribbie, R. A., & Flora, D. B. (2019). Evaluating equivalence testing methods for measurement invariance. *Multivariate Behavioral Research, 55*(2), 312–328. https://doi.org/10.1080/00273171.2019.1633617

Domínguez Castillo, J. G., Cisneros Cohernour, E. J., y Quiñonez Pech, S. H. (2025). Desafíos para la educación superior a distancia durante la pandemia por COVID-19. *Revista Educación, 49*(1), 1–20. https://doi.org/10.15517/revedu.v49i1.60880

Domínguez-Castillo, J. G., Cisneros-Cohernour, E. J., Quiñonez-Pech, S. H., y Ortega Maldonado, Á. (2024). Desafíos y barreras percibidas por las mujeres universitarias en sus cursos a distancia. *Revista científica electrónica de Educación y Comunicación en la Sociedad del Conocimiento, 24*(2), 409–433. https://doi.org/10.30827/eticanet.v24i2.31079

Dönmez, I. (2023). Breaking gender stereotypes: how interacting with stem professionals changed female students' perceptions. *Journal of baltic science education, 22*(6), 974-990. https://doi.org/10.33225/jbse/23.22.974.

Dost, G. (2024). Students' perspectives on the 'STEM belonging' concept at A-level, undergraduate, and postgraduate levels: an examination of gender and ethnicity in student descriptions. *International Journal of STEM Education, 11*(12). https://doi.org/10.1186/s40594-024-00472-9

Esteban-Ramiro, B., Moreno-López, R. y Marí., Y. R. (2024). ¿De ciencias o de letras? Brechas de género en el espacio educativo STEM [Science or Arts? Gender gaps in the STEM education space]. *European Public y Social Innovation Review, 9*, 1–20. https://doi.org/10.31637/epsir2024-1462

Fan, Z. G., Shi, X. L., Luo, Y., Chen, H. Y., & Wen, H.J. (2025). The Chinese version of the stigma of loneliness scale in people with chronic diseases: an assessment of psychometric characteristics. *BMC Public Health, 25*, 1–14. https://doi.org/10.1186/s12889-025-22743-y

Felicetti, V. L., & Cabrera, A. F. (2022). Students' experiences with graduate education in Brazil. A Confirmatory Factor Analysis Approach. *Revista de Investigación Educativa, 40*(2), 319–339. https://doi.org/10.6018/rie.513171

Gallego, L., & Casadiego, M. (2023). A Gender Gap Analysis on Academic Performance in Engineering Students on Admission and Exit Standardized Tests. I*ngeniería e Investigación, 43*(3). http://doi.org/10.15446/ing.investig.103276

González-Cid, C. (2024). Inequidades de género en STEM: El caso de académicas universitarias en Chile. *Cadernos de Pesquisa, 54*. https://doi.org/10.1590/1980531410649

González-Rogado, A. B., Ramos-Gavilán, A. B., Rodríguez-Esteban, M. A., y García-Holgado, A. (2025). Perception Disparity Between Women and Men on the Gender Gap in STEM at a Spanish University. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, 20*, 1-11. DOI: 10.1109/RITA.2024.3522255

Gregor, M. A., Weigold, I. K., Martin-Wagar, C. A., & Campbell-Halfaker, D. (2021). Tenure Expectations and Career Aspirations Among Female Assistant Professors in STEM. *Journal of Career Development, 49*(4), 890-905. https://doi.org/10.1177/08948453211005032.

Guilarte, V., Benarroch-Benarroch, A., & Enrique-Mirón, C. (2025). Psychometric Analysis of a Scale to Assess Education Degree Students' Satisfaction with Their Studies. *Education Sciences, 15*(6), 1-17. https://doi.org/10.3390/educsci15060660

Hair, J.F., Black, W.C., Babin, B.J., y Anderson, R.E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Jin, Y., & Chen, J. (2024). Regularized Variational Approximation for Partially Confirmatory Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 32*(3), 437–449. https://doi.org/10.1080/10705511.2024.2432612

Keller, L., Preckel, F., Eccles, J. S., & Brunner, M. (2022). Top-performing math students in 82 countries: An integrative data analysis of gender differences in achievement, achievement profiles, and achievement motivation. *Journal of Educational Psychology, 114*(5), 966–991. https://doi.org/10.1037/edu0000685.

Kinyota, M. (2021). A Portrait of the Gender Gap in STEM: A Focus on Identity Formation Among Final-year Undergraduate Students in Tanzania. *Journal of Education, Humanities and Sciences, 10*(3), 1–18.

Lendínez-Turón, A., Domínguez-Valerio, C. M., Orgaz-Agüera, F., & Moral-Cuadra, S. (2023). Public administration education towards Sustainable Development Goals: Psychometric analysis of a scale. *International Journal of Sustainability in Higher Education, 24*(6), 1177–1196. https://doi.org/10.1108/IJSHE-05-2022-0162

Lim, J. H., Wang, Y., Wu, T., Li, Z., & Sun, T. (2021). Walking on Gender Tightrope With Multiple Marginalities: Asian International Female Students in STEM Graduate Programs. *Journal of International Students, 11*(3), 647-665. https://doi.org/10.32674/jis.v11i3.2132

Lucena-Rodríguez, C., Invernón-Gómez, A. I., Ortiz-Marcos, J. M., & Sánchez-Mendías, J. (2025). Preparing future teachers for inclusive practices and disability: A systematic literature review. *International Journal of Instruction, 18*(2), 59-78. https://doi.org/10.29333/iji.2025.1834a

Marchionni, M., Gasparini, L., y Edo, M. (2018). *Brechas de género en América Latina: Un estado de situación*. Banco de Desarrollo de América Latina (CAF). https://scioteca.caf.com/handle/123456789/1306

Maxim, P. (2002). *Métodos cuantitativos aplicados a las ciencias sociales.* Oxford.

Mineiro, K.M.L., Laurett, R., Nossa, V., & Nossa, S.N. (2025). Management accountant skills: An exploratory factor análisis. *The International Journal of Management Education, 23*(3), 1-13. https://doi.org/10.1016/j.ijme.2025.101209

Morata-Ramírez, M.ª A., Holgado-Tello, F. P., Barbero-García, I., & Méndez, G. (2015). Análisis factorial confirmatorio: recomendaciones sobre mínimos cuadrados no ponderados basadas en el error tipo I de Chi-cuadrado y RMSEA. *Psychological Action, 12*(1), 79-90. https://dx.doi.org/doi.org/10.5944/ap.12.1.14362

Naciones Unidas. (2023). *Objetivos y metas de desarrollo sostenible*. https://www.un.org/sustainabledevelopment/es/sustainable-development-goals/

Neri, C., Torres, S., Fajardo, N., Retamoza, P., Rodríguez, V., Castillo, C., y Pérez, L. (2022). Análisis del comportamiento en procesos de ingreso y egreso de mujeres en la carrera de Ingeniería Química en el Centro Universitario de Ciencias Exactas e Ingenierías de la Universidad de Guadalajara (México). *Revista Iberoamericana de polímeros, 23*(4), 119–125.

O'Connor, P. (2023). Is gendered power irrelevant in higher educational institutions? Understanding the persistence of gender inequality. *Interdisciplinary Science Reviews, 48*(4), 669-686. https://doi.org/10.1080/03080188.2023.2253667.

OECD. (2022). *Education at a glance 2022: Who graduates from tertiary education? OECD*. https://www.oecd.org/education/education-at-a-glance/(accessed on 04/02/2025).

ONU Mujeres. (2025). *Brechas de género en STEM. Una mirada al sector energético*. https://lac.unwomen.org/sites/default/files/2025-03/informe-stem-onumujeres_final-comprimido.pdf

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2024a). *Changing the equation: securing STEM futures for women.* https://unesdoc.unesco.org/ark:/48223/pf0000391384

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2024b). *La UNESCO en acción por la igualdad de género: 2022-2023*. https://unesdoc.unesco.org/ark:/48223/pf0000387300_spa

Paige, S.R., Flood-Grady, E., Krieger, J.L., Stellefson, M., & Miller, M.D. (2019). Measuring health information seeking challenges in chronic disease: A psychometric analysis of a brief scale. *Chronic Illness,17*(2), 151-156. https://doi.org/10.1177/1742395319836476

Papagiannopoulou, T., Vaiopoulou, J., & Stamovlasis, D. (2023). Teachers' Readiness to Implement STEM Education: Psychometric Properties of TRi-STEM Scale and Measurement Invariance across Individual Characteristics of Greek In-Service Teachers. *Education Sciences, 13*(3), 1-13. https://doi.org/10.3390/educsci13030299

Piva, E., & Rovelli, P. (2022). Mind the gender gap: the impact of university education on the entrepreneurial entry of female and male STEM graduates. *Small business economics, 59*(1), 143–161. https://doi.org/10.1007/s11187-021-00525-1.

Pownall, M., y Heflick, N. (2024). Male psychologists and female mathematicians: Gender beliefs and undergraduate degree choices. *Journal of Community & Applied Social Psychology, 34*(2), 1-16. https://doi.org/10.1002/casp.2784

Salahshouri, A., & Mohamadian H. (2024). Validation of the Dundee Ready Education Environment Measure in Iran through factor analysis. *Education for health, 37*, 138-146. https://doi.org/10.62694/efh.2024.9

Shahin, M., Ilic, O., Gonsalvez, C., & Whittle, J. (2021). The impact of a STEM-based entrepreneurship program on the entrepreneurial intention of secondary school female students. *International Entrepreneurship and Management Journal*, 17, 1867–1898. https://doi.org/10.1007/s11365-020-00713-7

Shanks, D.R., Coles, H.A., & Yeo, N. (2024). A re-evaluation of gender bias inreceptiveness to scientific evidence of gender bias. *Royal society open science, 11*(9), 1-15. https://doi.org/10.1098/rsos.240419.

Slattery, O., Prendergast, M., y Riordáin, M. (2023). Navigating a male dominated domain: experiences of female STEM students in higher education in Ireland. *Irish Educational Studies, 42*(4), 861-880. https://doi.org/10.1080/03323315.2023.2261418

Stewart, R., Wright, B., Smith, L., Roberts, S., y Russell, N. (2021). Gendered stereotypes and norms: Asystematic review of interventions designed to shift attitudes and behaviour. *Heliyon, 7*(4), 1-15. https://doi.org/10.1016/j.heliyon.2021.e06660

Tomé-Fernández, M., Ortiz-Marcos, J. M., & Olmedo-Moreno, E. M. (2019). Educational Environments with Cultural and Religious Diversity: Psychometric Analysis of the Cyberbullying Scale. *Religions, 10*(7), 1-16. https://doi.org/10.3390/rel10070443

Vargas Jiménez, M., Domínguez Castillo, J. G., & Cisneros-Cohernour, E. (2025). Retos de género en STEM: Un enfoque en comunidades rurales. En J. M. Ortiz-Marcos & M. Tomé-Fernández (Coords.), *Transferencias educativas para la convivencia y la paz social en contextos interculturales* (pp. 107–126). Editorial Universitas.

Verdugo-Castro, S., García-Holgado, A., Sánchez-Gómez, M. C., y García-Peñalvo, F. J. (2025). Modelos y referentes que influyen en los Estereotipos de Género en STEM: un estudio de caso en España. *Revista de Investigación Educativa,* 43, 1-27. https://doi.org/10.6018/rie.597191

World Bank. (2024, 07 de octubre). *Estrategia de Género del Grupo Banco Mundial para 2024-2030: Acelerar la Igualdad de Género para Poner Fin a la Pobreza en un Planeta Habitable (Spanish)*. Washington D. C.: World Bank Group. http://documents.worldbank.org/curated/en/099748510072434726

World Economic Forum (WEF). (2025, 11 de junio). *Global Gender Gap Report 2025*. https://www.weforum.org/publications/global-gender-gap-report-2025/

Yamamoto, S. H. (2023). Preliminary Psychometric Assessment of STEM Attitude Measure for U.S. High School Students With Disabilities. *Journal of Psychoeducational Assessment, 42*(3), 349-356. https://doi.org/10.1177/07342829231218767

Yang, D., & Chittoori B. (2022). Investigating Title I school student STEM attitudes and experience in an after-school problem-based bridge building project. *Journal of STEM Education, 23*(1), 17–24. https://www.jstem.org/jstem/index.php/JSTEM/article/view/2504

Yang, X., & Gao, C. (2021). Missing Women in STEM in China: an Empirical Study from the Viewpoint of Achievement Motivation and Gender Socialization. *Research in Science Education, 51*, 1705-1723. https://doi.org/10.1007/s11165-019-9833-0.

Yokoyama, H. M., Ikkatai, Y., McKay, E., Inoue, A., Minamizaki, A., & Kano, K. (2024). Can affirmative action overcome STEM gender inequality in Japan? Expectations and concerns. *Asia Pacific Business Review, 30*(3), 543–559. https://doi.org/10.1080/13602381.2024.2320547.