Gender, self-efficacy and performance in a mathematics test: The moderating role of the educational center

Género, autoeficacia y desempeño en una prueba de matemática: El papel moderador del centro educativo

Gênero, auto-eficácia e desempenho em um teste matemático: O papel moderador da escola

José Andrey Zamora-Araya[1]*, **Eiliana Montero-Rojas**[2], **Vanessa Smith-Castro**[3],
Tania Elena Moreira-Mora[4], **Pablo Zamora-Calvo**[5], **Karla Quintero-Arias**[5], **Stefani Matarrita-Muñoz**[5]

Abstract

[**Objective**] The purpose of the present study was to analyze the relationship between students' mathematical self-efficacy and performance in the national high school graduation mathematics test, taking into account other relevant psychosocial predictors and the moderating effects of sex and educational center. [**Methodology**] The design of the study was observational and cross-sectional, with the participation of 487 eleventh-grade students from public secondary schools in the Greater Metropolitan Area of Costa Rica. Several instruments were applied for measuring hostile and benevolent sexism, perceptions of gender equality in mathematical abilities, mathematical self-efficacy and reasoning abilities. [**Results**] The results of the multilevel Bayesian random effects analysis confirmed that scores were higher for men than for women in measurements of self-efficacy, where the sex of the student body was a moderator in the relationship of mathematical self-efficacy and test scores. In addition, it was found that, depending on the sex of subjects, the association between self-efficacy and performance was also moderated by the educational center. [**Conclusions**] Among the conclusions, it is especially noteworthy that if the moderating effect of the educational center between self-efficacy and performance differs by sex, this implies that actions aimed at promoting self-efficacy will not have the same result in men and women, showing the need to develop differentiated interventions.

* Corresponding author

José Andrey Zamora-Araya, jzamo@una.ac.cr, http://orcid.org/0000-0001-6050-5850
Eiliana Montero-Rojas, eiliana.montero@ucr.ac.cr, http://orcid.org/0000-0002-6770-792X
Vanessa Smith-Castro, vanessa.smith@ucr.ac.cr, http://orcid.org/0000-0001-6348-4223
Tania Elena Moreira-Mora, tmoreira@itcr.ac.cr, https://orcid.org/0000-0002-8955-0804
Pablo Zamora-Calvo, pablo.zamoracalvo@gmail.com, https://orcid.org/0000-0001-7857-1311
Karla Quintero-Arias, karla.quinterocr@gmail.com, https://orcid.org/0000-0001-7088-7178
Stefani Matarrita-Muñoz, stefmatarrita@hotmail.com, https://orcid.org/0000-0002-1222-1981

1   Department of Mathematics, Universidad Nacional, Heredia, Costa Rica.
2   Institute for Psychological Research, Department of Statistics, Universidad de Costa Rica, San José, Costa Rica.
3   Institute for Psychological Research, Universidad de Costa Rica, San José, Costa Rica.
4   Department of Orientation and Psychology, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica
5   Universidad de Costa Rica, San José, Costa Rica.

Keywords: Mathematical self-efficacy; mathematics performance; Bayesian multilevel models; sexism.

Resumen 🔊

[Objetivo] El propósito del estudio fue examinar la relación entre autoeficacia matemática y el rendimiento en la prueba nacional de bachillerato en matemáticas, que considera otros predictores psicosociales relevantes y los efectos moderadores del sexo y del centro educativo. **[Metodología]** El diseño del estudio fue de corte observacional y transversal, con la participación de 487 estudiantes de undécimo año, provenientes de colegios públicos de la Gran Área Metropolitana de Costa Rica. Se aplicaron varios instrumentos para medir: sexismo hostil y benevolente, percepciones de igualdad de género en las habilidades matemáticas, autoeficacia matemática y habilidades de razonamiento. **[Resultados]** Los resultados del análisis multinivel bayesiano de efectos aleatorios confirmaron las diferencias a favor de los hombres en la medida de autoeficacia, donde el sexo del estudiantado fue un moderador entre la relación de la autoeficacia matemática y el puntaje de la prueba. Además, se encontró que, según el sexo, la asociación entre autoeficacia y rendimiento también se encontró moderada por el centro educativo. **[Conclusiones]** Entre las conclusiones se destaca que, si el efecto moderador del colegio entre autoeficacia y rendimiento difiere por sexo, esto implica que las acciones dirigidas a promover la autoeficacia no tendrán el mismo resultado en los hombres y en las mujeres, lo que evidencia la necesidad de desarrollar intervenciones diferenciadas.

Palabras clave: autoeficacia matemática; desempeño matemático; modelos multinivel bayesianos; sexismos

Resumo 🔊

[Objetivo] O objetivo do estudo foi examinar a relação entre auto-eficácia matemática e desempenho no teste nacional de bacharelado em matemática, levando em conta outros preditores psicossociais relevantes e os efeitos moderadores relacionados ao sexo e à escola. [**Metodologia**] O desenho do estudo foi de coorte observacional e transversal, com a participação de 487 estudantes do décimo primeiro ano das escolas públicas da Grande Área Metropolitana da Costa Rica. Vários instrumentos foram aplicados para medir: sexismo hostil e benevolente, percepções de igualdade de gênero em habilidades matemáticas, auto-eficácia matemática e capacidade de raciocínio. **[Resultados]** Os resultados da análise multinível Bayesiana de efeitos aleatórios confirmaram diferenças a favor dos homens na medida da auto-eficácia, onde o sexo dos estudantes foi um moderador entre a relação de auto-eficácia matemática e a pontuação no teste. Além disso, encontrou-se que, dependendo do sexo, a associação entre auto-eficácia e desempenho também foi moderada pela escola. **[Conclusões]** Entre as conclusões, destaca-se que, se o efeito moderador da escola entre auto-eficácia e desempenho difere por sexo, isso implica que as ações destinadas a promover a auto-eficácia não terão o mesmo resultado em homens e mulheres, o que evidencia a necessidade de desenvolver intervenções diferenciadas.

Palavras-chave: autoeficácia matemática; desempenho matemático; modelos Bayesianos multiníveis; sexismos.

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

Introduction

During the last 30 years, a number of studies have indicated that women have stronger abilities than men in the areas of reading and writing, and that men have stronger abilities than women in the areas of mathematics and science (Else-Quest, Hyde & Linn, 2010; Hyde, Lindberg, Linn, Ellis y Williams, 2008; Lindberg, Hyde, Peterse, y Linn; 2010; Lynn y Mikk, 2009; Montero Rojas, Moreira-Mora, Zamora-Araya y Smith-Castro, 2021; Moreira Mora, Zamora-Araya, Smith Castro y Montero Rojas, 2021; Voyer y Voyer, 2014). Based on analysis of the results of the PISA standardized tests, the most recent studies have found that women tend to outperform men in reading by more than 25 points (OECD, 2014), while men's performance in the science test is, on average, four (4) points higher than that of women, although in some countries these differences between the sexes in science results can be as much as 15 points (OECD, 2015).

As a result of these disparities, there is a strong interest in the scientific and educational community in identifying those socio-structural, cultural, educational and psychological factors that contribute to maintaining these gaps. Among psychological factors, the implication of mathematical self-efficacy stands out as an important predictor, both of academic performance and of motivation and persistence in careers that demand mathematical skills (Jordan & Carden, 2017). Likewise, there is empirical evidence demonstrating that women show less academic confidence than their male counterparts, particularly in mathematical contexts (Huang, 2013), and that these differences are linked to cultural stereotypes and sexist beliefs that have an impact on the perception students have about their abilities and skills (Smith-Castro, Montero-Rojas, Moreira-Mora & Zamora-Araya, 2019).

However, research also strongly indicates that individual differences in skills and attitudes are developed, maintained and strengthened depending on the immediate educational contexts of students. Studies of factors associated with academic performance have shown that the characteristics of educational centers (including factors such as location, schedule and school climate) also play an important role in determining student performance (Murillo Torrecilla & Román Carrasco, 2009; Martínez et al., 2017).

Given the multidimensional and multifactorial nature of academic performance, this article aims to contribute to the study of the effect of mathematical self-efficacy on the mathematical performance of young people who are in their eleventh grade in public schools in the Greater Metropolitan Area (GAM) of Costa Rica, given the moderating role that students' gender and the educational center they attend play in this relationship. To do so, a Bayesian multilevel model technique was used (Kaplan, 2014), which has analytical advantages over the frequentist models used in previous research (Montero-Rojas et al., 2021). This approach can contribute to the analysis of gender and performance gaps in mathematical contexts that are observed in Costa Rica, by using more robust tools to identify the origin of these inequities.

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

Theoretical framework

Self-efficacy and academic performance

The concept of self-efficacy was introduced by Bandura (1977) as a central aspect of his social cognitive theory to describe the degree to which people consider themselves to be agents with effective control over their actions. In general, it is concerned with the degree of people's convictions about their effectiveness when dealing with the contingencies of daily life (Schwarzer, Bäßler, Kwatek, Schröder, & Zhang, 1997). A significant number of studies have shown that high perceived self-efficacy increases the probability of academic success by virtue of its association with motivation and learning (Zamora, 2020). In particular, it has been observed that students with a high sense of mathematical self-efficacy tend to participate more in challenging tasks, invest more effort in their execution, are more persistent and show better academic performance compared to those who lack such confidence (Bong, 2004).

As mentioned previously, there is evidence that men and women differ in their levels of self-efficacy. A meta-analytic investigation of 187 studies, with a total of 68,429 participants (Huang, 2013) identified differences with low effects in favor of men (d = 0.08). Moderator analysis showed that the degree of the effect varied by academic domain or area: women showed higher artistic and linguistic self-efficacy than men, while men showed higher mathematical and computational self-efficacy than women. Furthermore, the results showed that gender differences in academic self-efficacy also varied with age. The largest effect was observed among respondents at the end of adolescence. These trends have also been found in more recent studies (Navarro, Flores & González, 2019; Recber, Isiksal & Koç, 2018).

Sexist beliefs and self-efficacy

According to several theoretical models (Glick & Fiske, 1996; 2011; Spencer, Steele & Quinn, 1999; Steele, Spencer & Aronson, 2002), differences in self-efficacy between men and women have multiple causes, but are undoubtedly linked to cultural factors such as gender stereotypes and sexist beliefs that persist even today.

Sexist beliefs may be divided into two broad categories: hostile and benevolent. The first is based on the false belief that characterizes women as a subordinate and inferior group, legitimizes social control exercised by men over women, distrusts women and penalizes those who break with traditional roles. Benevolent sexism, on the other hand, is based on the false belief that idealizes women as wives, mothers and romantic objects, who must be protected by men because of their delicacy and inability to function independently (Glick & Fiske 1996).

These beliefs can negatively affect women's performance through multiple mechanisms, most notably through effects on self-esteem, self-concept, self-efficacy, motivation and perseverance, which, in turn, affect performance in different aspects of women's academic, work, and social lives (Guiso, Monte, Sapienza and Zingales, 2008). For example, it has been shown that women's performance is lower (compared to men) when negative beliefs about their abilities are activated in contexts involving mathematics (Spencer, Steele, & Quinn, 1999). Research has further shown that the negative impact of gender stereotypes on

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

women's math performance can have effects of up to 20 points on the math component of standardized tests (Walton & Spencer, 2009).

On the other hand, recent data from Zell et al. (2016) have shown that people who support sexist beliefs tend to perceive more gender differences and exaggerate the extent of such differences in several domains, including math tests, academic skills, and interest in scientific fields. This exaggeration of differences can affect students' self-concepts, which can harm their performance.

Finally, it has been shown that sexist beliefs and gender stereotypes can affect academic performance through the attitudes of socializing agents such as peers, parents, and teachers (Jacobs & Eccles, 1992). Specifically, evidence shows that girls whose mathematical abilities are repeatedly challenged by their environment develop a lower mathematical self-concept, less confidence in their mathematical aptitude, and are less motivated than their male peers (Leaper & Brown, 2014).

The moderating role of educational centers

Research indicates that individual differences are developed, maintained and strengthened in specific educational contexts. Studies on factors associated with academic performance have shown that the characteristics of educational centers such as location, schedule and school climate play a fundamental role in student performance (Murillo & Román, 2009).

As a result, academic performance can be affected by other factors such as socioeconomic level and the educational climate of educational centers. In this regard, Berkowitz et al. (2015) pointed out that the

academic climate can be associated with: (a) compensation, especially for low-income students, since it contributes positively to achievement; (b) mediation, since the socioeconomic level of the educational center has an influence on the academic climate and thus on performance, together with other factors such as commitment to the institution; and (c) moderation, since the relationship between academic performance, the socioeconomic level of the educational center and the student body can be moderated by the academic climate of the educational center. Likewise, a strong educational environment not only promotes better academic performance, but can even help counteract the influence of negative factors related to conditions of poverty and problems at home (O' Mally et al., 2015).

In addition, the socioeconomic and demographic conditions of the areas in which the educational centers are located, may in many cases determine the quality of the education their students receive. In this sense, the academic climate in the educational center plays a moderating role between socioeconomic level and academic performance, to the point that institutions with a favorable academic climate can help reduce the gap in academic performance, in spite of differences in socioeconomic backgrounds (Berkowitz, 2021; Ruiz, McMahon and Jason, 2018).

On the other hand, a previous study confirmed that there are differences in the performance of men and women in mathematical context tests in Costa Rica, which were mainly predicted by mathematical self-efficacy (Montero-Rojas et al., 2021). This study further showed that this association was also moderated by the educational center – that is, there was evidence that confirmed that the magnitude of the

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

relationship between self-efficacy and the score in the mathematical context tests varied according to the educational center. In statistical terms, an effect of the educational center was identified in the relationship between mathematical self-efficacy and test scores. However, the analytical tool used in this research, multilevel frequentist models, does not allow direct estimation of the magnitude of the effects of the schools' academic climates. The present analysis proposes to overcome this limitation by applying multilevel Bayesian models.

At the international level, there is abundant research that relates various socio-emotional and attitudinal dimensions with indicators of academic performance, but those that explicitly study the moderating effects of educational centers on these relationships are less common. Among these, the research by Chamizo-Nieto *et al.* (2021) is notable for analyzing how the quality of the student-teacher relationship moderates the influence of emotional intelligence on academic performance in a sample of Spanish adolescents; based on its results, the authors recommended that it is crucial to promote better relationships between students and teachers, especially when the former show low levels of emotional intelligence. This study also confirmed that "flourishing" is a mediating variable in the model, defined as a pathway for growth and resilience.

Likewise, the study by Mella *et al.* (2021) explores how affective factors of school adjustment influence the relationship between socioemotional competencies and school grades in language and mathematics, obtained through a network approach, in a sample of 3,400 French secondary education students. The results showed that self-regulation in the educational center had the greatest weight in the network. In particular, this variable measured the relationships between interpersonal socioemotional competencies and grades.

Advantages of the Bayesian approach over the frequentist approach

In general terms, the Bayesian approach represents a paradigm shift with respect to the frequentist approach, which is partly evidenced in the methodology used to obtain estimates. While the Bayesian approach in most cases uses Markov chains combined with the Monte Carlo method (MCMC), the frequentist approach uses maximum likelihood or restricted maximum likelihood estimates, with their respective assumptions about data distribution (Finch *et al.*, 2014).

Another difference of the Bayesian approach is the use of prior distributions, which represent the extent of knowledge about the parameter to be estimated, which are not present in the frequentist approach and which, together with the sample data, are used to estimate the parameter. As Ruiz-Benito (2018, p. 136) points out, it consists of: "an iterative learning process in which conclusions are reached about a phenomenon (a posteriori probability) based on prior knowledge about the system (a priori probability) and of new evidence (information from the data)."

In the case of multilevel models, according to Finch *et al.* (2014) there are several advantages to using a Bayesian model instead of a frequentist one. These include: (1) Bayesian models are not based on assumptions about data distribution; (2) when the model to be estimated is very complex, the frequentist approaches may not converge; (3) for small samples, the Bayesian approach could provide more accurate

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

estimates of the model parameters; and (4) in frequentist models that perform well, Bayesian estimates are still correct, and in these cases the results tend to be similar.

In brief, given the importance of incorporating the effects of educational contexts into the analysis, and recognizing the potential of Bayesian models and their infrequent use in the context being analyzed, this study aims to:

1) Determine if the relationship between self-efficacy and performance on a standardized math test is conditioned by students' gender.
2) Determine the existence of a moderating effect of the educational center on the relationship between mathematical self-efficacy and performance on such a test; and
3) Estimate the moderating effects of educational centers on the relationship between self-efficacy and test performance.

As will be seen in following sections of this document, all the analyses that made it possible to meet these objectives were carried out using the variables of sexism and gender stereotypes. As indicated in the background discussion, these variables represent important predictors of performance and merit being included in analytical models.

Methodology

The study is observational and cross-sectional. The sample consisted of 487 eleventh-grade students, of which 262 were women. The 10 participating educational centers were randomly selected from public academic secondary schools in the Greater Metropolitan Area (GAM) in 2015.

Instruments

Hostile sexism and benevolent sexism. The Spanish version of the ambivalent sexism inventory (Expósito, Moya, & Glick, 1998) was used. The instrument contains 22 items that measure sexist antipathy towards women, based on the perception that women seek control over men (hostility dimension) and the belief that women are delicate creatures who should be confined to traditional roles (benevolence dimension). Participants marked each item on a five-point Likert scale from 1 (completely disagree) to 5 (completely agree).

Perceptions of gender equality in mathematical skills. These are 4 items taken from the scale of Mathematics as a neutral domain (Forgasz, Leder, & Gardner, 1999), which assesses the extent to which participants believe that women are as good as men in mathematics. As in the previous scale, participants evaluated each item on a five-point Likert scale.

Mathematical self-efficacy. These are 12 Likert-scale items taken from the subscale of self-confidence in learning mathematics (Fennema & Sherman, 1976) and are used to measure the extent to which each respondent is confident in their ability to learn and successfully perform mathematical types of tasks, with the same response scale of the previous measures.

Reasoning skills. Abstract reasoning was measured by a subset of items from the Figures Reasoning Test (PRF) (Montero, *et al.*, 2013). These 17 items of seriation tasks, with six options each, are indicators of fluid cognitive functioning, where speed is one of the evaluation components. Participants received one point for each correct answer.

Performance in the standardized national high school graduation mathematics test. The percentage of correct answers in

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora •
Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

7

the high school mathematics test was used. It is a 60-item single-selection exam that covers concepts and procedures in mathematical competence, subdivided according to the general abilities emphasized in the national Mathematics study programs (Mena, 2015).

Data collection protocol

After approval of the University of Costa Rica's bioethics committee was obtained, permissions were requested from the educational centers. Once permissions were granted, students were invited to participate voluntarily in the study. In the classrooms of their respective schools, they completed the reasoning test and the attitudinal scales. The scores of the participants in the high school graduation mathematics test were collected using the databases of the Quality Management and Evaluation Office of the Ministry of Public Education of Costa Rica (DGEC), with the approval of the students through their informed consent.

Statistical analysis

The analysis procedure involved the estimation of multilevel random effects models for the coefficients that connect mathematical self-efficacy with test grades. Both frequentist and Bayesian approaches were used for this estimation; the samples for men and women were analyzed separately (multigroup analysis), and were controlled for basic reasoning abilities and other attitudinal variables.

A Bayesian multilevel analysis was used in which a sample of n students is taken and there is information for each student i attending the educational center j, j=1, 2, …,10, about the grade, on a 0-100 scale in the 2014 high school mathematics graduation test that was used as the dependent variable (y_{ij}), and six predictor or independent variables described below:

x_1: Age in completed years
x_2: Percentage of correct answers in the reasoning test with figures
x_3: Participant's average score on the benevolent sexism scale
x_4: Participant's average score on the hostile sexism scale
x_5: Participant's average score on the mathematics equity scale
x_6: Participant's average score on the mathematics self-efficacy scale

In the final 4 attitudinal scales ($x_3 - x_6$), the average score varies on a scale that ranges from 1 to 5. In the Bayesian model, x_6 functions as a random slope that varies from one educational center to another. Therefore, the specification of the model is as follows:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \beta_5 x_{5ij} + \beta_6 x_{6ij} + u_{0j} + x_{6ij} u_{1j} + e_{ij} \quad (1)$$

The terms in the random part of the model are: u_{1j} which represents the random coefficient of the self-efficacy measure, u_{0j} the random intercept, e_{ij} the level 1 residual and β_0, β_1, ..., β_6 which represent the intercept and coefficients of the fixed part of the model. The model also assumes that

$$u_{0j} \sim N(0, \sigma_{u0}{}^2) \, , \, u_{1j} \sim N(0, \sigma_{u1}{}^2) \text{ y} \atop e_{ij}{}^2 \sim Inv \, Gamma(0.01, 0.01) \quad (2)$$

Non-informative prior distributions are used, which represent the lack of knowledge about the parameters to be assessed. Stata 15.1 software was used to estimate

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

the model, which, by default, establishes non-informative priors for the model parameters, which for the proposed case are the following:

$$\beta_0, \beta_1, \ldots, \beta_6 \sim N(0,10000),$$
$$\sigma_{u0}{}^2 \sim Inv\ Gamma(0.01,0.01),$$
$$\sigma_{u1}{}^2 \sim Inv\ Gamma(0.01,0.01) \quad (3)$$

That is, for coefficients of independent variables and the intercept, the normal distribution was used as prior and the inverse gamma distribution for the variances. The use of prior distributions, whether fuzzy or informative, is one of the main differences between Bayesian and frequentist estimates.

In the case of multilevel models, one of the main advantages of Bayesian estimation over frequentist estimation is that it allows direct estimation of the random effects for each of the clusters, in this case, educational centers, while in the frequentist models it would be done through the matrix of variances and covariances of the random effects. This characteristic of the Bayesian estimation strengthens the results, since it is generally of interest to not only weigh whether there are important moderating effects due to the clusters, but also, as in the present case, to estimate

those effects in each of the conglomerates to support substantive interpretations and decision making (Montero, Rojas, Negrín, & Francis, 2015).

Analysis and results

At a general level, it is observed that the basic results are similar for both estimation methods, evidence that points out to the compliance with the frequentist assumptions and the use of non-informative prior distributions in the Bayesian approach. Regarding the stability of estimates using the Bayesian model, there is convergence in the model's estimates (see Figure 1, Annex 1).

Regarding the first objective, it was confirmed, in accordance with the results of previous research, that average self-efficacy was higher in men than in women. Specifically, the effect size of the difference in mathematical self-efficacy, according to sex, was 47.98%; this is defined as the difference of the means divided by the combined standard deviation for both groups, according to the following formula:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1{}^2 + S_2{}^2}{2}}} \quad (4)$$

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

Table 1.
Fixed and random effects estimates using the Bayesian approach

| Predictors | High school graduation mathematics exam | | | |
| | Women | | Men | |
	B	**Beta**	**B**	**Beta**
Fixed effects				
Age	-2.822	-0.172	-2.582	-0.148
Correct Reasoning	0.139	0.210	0.270	0.393
Benevolent Sexism	-1.604	-0.087	-1.038	-0.061
Hostile Sexism	-1.311	-0.067	-0.800	-0.045
Gender equality	1.487	0.066	0.087	0.004
Mathematical self-efficacy	3.143	0.189	6.089	0.400
Intercept	89.570	-0.139	70.391	-0.019
Random Effects (SD)	Estimated	Standardized	Estimated	Standardized
Self-efficacy	0.975	0.164	1.137	0.267
Constant	3.229	0.340	4.585	0.399
Residual	11.144	0.796	10.813	0.767
Rho	8 %	18 %	16 %	28 %

Source: Calculated by the authors during this investigation.

In terms of the relationship between mathematical self-efficacy and the grade in the high school test, according to Table 1, the standardized fixed coefficient for women is 0.19 and for men it is 0.40, i.e., almost double, which confirms the moderating effect of sex. In both models, the control variables that have the highest regression coefficients are age (inverse relationship) and the reasoning test score (direct relationship), while the relationship with the other attitudinal indicators was marginal.

Regarding the second objective of determining the existence of a moderating effect of the educational center, the magnitude of the random coefficients for the relationship between self-efficacy and test grade is presented in Table 1, according to sex. Standardized random effects values greater than 0.10 were obtained for both the constant and for self-efficacy. Specifically, the random coefficients for estimating the constant or intercept are 0.34 for women and 0.40 for men, respectively, and for self-efficacy they were 0.16 for the former and 0.27 for the latter. On the one hand, the random effects for the constant account for the starting differences in student performance which, according to these data, were quite substantial, despite the fact that all of the educational centers involved are public secondary day schools in the GAM. On the other hand, the random effects for self-efficacy indicate that the relationship between this variable and student grades can vary for both men and women depending on the schools that they attend.

In the case of the third objective, the moderating effects of the educational center on the relationship between self-efficacy and mathematical performance were explored in more detail by considering the samples of men and women separately, given the results obtained when achieving the first objective. To do so, the estimation of the random effects of self-efficacy was directly obtained for each educational center using Bayesian estimation/approach, which in the frequentist case can

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

only be approximated based on values predicted by the same model. The results of this analysis are presented in Table 2. The educational centers are identified by numbers to protect the confidentiality of the participating institutions.

It should be noted that the random effects for each educational center were calculated taking as a reference the value of the fixed effect. For example, the random effect of -0.424 for women in educational center 3 indicates that the estimated value is lower by 0.424 than the value of the fixed coefficient of 3.143 units; its estimated value is therefore 2.719, as indicated in the column labeled Estimation, in Table 2.

Based on these results, the following categorization of educational centers can be proposed:

1) Those educational centers in which the moderating effect is negative for both sexes; that is, the direct relationship between mathematical self-efficacy and performance on the mathematics test is diminished, decreased, or reduced in relation to the estimated fixed effect of 3.143 for women and 6.089 for men. In these educational centers, mathematical self-efficacy seems to have a lower predictive power for the grades obtained in the test, for both sexes. A reasonable hypothesis to explain this behavior, which should be evaluated in future research, is that there may be other variables that influence the relationship between mathematical self-efficacy and test performance.

2) Cases in which the random effect is positive for both sexes. In these schools, the results confirmed that higher levels of self-efficacy are associated with better performance on the test, as predicted by theory and previous evidence. Furthermore, in these educational centers the direct relationship between these two variables is stronger, compared to the magnitude of what we can call the "average effect" represented by the fixed coefficient in the model. It would therefore be important to know if the students' self-efficacy is promoted in these educational centers, since this would result in better performance. If this is

Table 2

Estimation of the moderating effects of educational centers on the relationship between mathematical self-efficacy and performance in the high school graduation mathematics test

Educational center code	Women			Men			Moderating effect of Educational Center
	n	Random effect	Estimate	n	Random effect	Estimate	
3	27	-0.424	2.719	28	-1.241	4.848	Negative for both
5	28	-2.356	0.787	32	-2.177	3.912	Negative for both
6	12	-1.626	1.517	7	-4.445	1.644	Negative for both
10	39	-0.409	2.734	38	-0.251	5.838	Negative for both
4	13	1.642	4.785	8	2.088	8.177	Positive for both
7	45	1.969	5.112	41	3.003	9.092	Positive for both
9	23	1.473	4.616	21	1.375	7.464	Positive for both
8	40	1.800	4.943	26	0.305	6.394	Positive for both
2	11	0.006	3.149	6	1.529	7.618	Positive for men, no effect for women
1	24	-0.458	2.685	18	1.034	7.123	Positive for men and negative for women

Source: Calculated by the authors during this investigation.

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

the case, appropriate practices could be identified to be replicated in other educational centers.

3) Educational centers in which the random effects vary strongly depending on students' sex, a result that should be studied in more depth. It could be hypothesized that if mathematical self-efficacy is promoted in these educational centers, this would have effects that would differ depending on whether the subject is a man or a woman.

It is important to emphasize that this interpretation must be considered with caution, given that the samples in the educational centers are relatively small, especially in the cases of educational centers 2, 4 and 6.

Conclusions

The object of the present study was to examine the relationship between mathematical self-efficacy and performance in the national high school graduation mathematics test, taking into account other relevant psychosocial predictors and the moderating effects of sex and educational center.

Our results confirm the existence of higher average values of scores for self-efficacy for men that for women, with an effect size of almost 50%, which is consistent with the findings discussed in the introductory sections of this document (Huang, 2013; Jordan & Carden, 2017; Smith-Castro, *et al*., 2019).

On the other hand, the results obtained also indicate that students' sex plays an important role as a moderator of the relationship between mathematical self-efficacy and test scores, although in both cases self-efficacy is a relevant predictor. Specifically, results of analyses carried out in this study indicate that if the student is a man, the magnitude of the average relationship between self-efficacy and achievement will be almost double the magnitude of the relationship when the person is a woman.

Likewise, the finding that the association between self-efficacy and performance according to sex is also moderated by the educational center is important. This finding could not have been detected using other statistical techniques. The magnitude of the moderating effect of the educational center varies, although in all cases it is positive, which once again confirms the importance of this finding, even when the level of basic reasoning skills is controlled for or kept constant. In other words, in a hypothetical case a specific student with a certain level of reasoning skills and a given level of mathematical self-efficacy is investigated, his or her grade in the mathematics test will be higher or lower depending on whether he or she is a male or female and the educational centers he or she attends. In some educational centers the relationship is positive and strong, while in others it is weaker. In other cases, the effect of self-efficacy on performance is higher for one of the sexes in the same educational center. In other words, if the educational center's moderating effect for the relationship between self-efficacy and performance varies by sex, this means, for example, that actions aimed at promoting self-efficacy will not produce the same result in mathematics performance in men and women. Results will be of higher magnitude in the group in which the relationship is stronger. This could indicate the need for differentiated interventions for men and women to promote mathematical self-efficacy in educational contexts.

Based on these conclusions, recommendations are proposed at the practical and investigative level. In practical terms,

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

12

it will be important for teachers to promote mathematical self-efficacy in all their students, but especially among women; this must be achieved through explicit policies of the Ministry of Public Education, clearly expressed as learning objectives in its programs of study. To ensure success of these interventions, it will be necessary to carry out diagnoses on beliefs, attitudes and teaching practices in terms of the supposed gender differences in mathematical contexts to guide their design and execution. Currently, training plans in the teaching of mathematics lack coverage of the issue of gender equality, not only as part of the transversal axes of the curriculum, but also as specific content, directly evaluable, throughout the formation of future mathematics teachers (UNESCO, 2020).

At the level of the research agenda, this initial study opens the possibility of further investigating the mechanisms through which educational centers exert moderating effects, rising to the level of research oriented towards the generation and falsification of hypotheses. It will also be important to collect information on characteristics of educational centers (including size, composition of the teaching body, level of specialization of the teaching staff, educational climate and location), to detect those attributes of the educational centers which influence the differences observed in this investigation. The estimates of the educational center's moderating effects derived here are descriptive; for future research it is therefore suggested that larger samples be selected. Likewise, it would be desirable to replicate the study in other educational centers and in different geographical areas, as our findings cannot be generalized to the great variety of educational centers that make up our educational system, and the diversity of educational experiences that occur in the different geographical areas in which they are located.

Funding

Informed consent

In accordance with the guidelines of the Scientific Ethics Committee of the University of Costa Rica, participants were informed of the objectives of the study, the confidentiality nature of the use of their data, and their freedom to stop participating at any time, through an informed consent form included in the questionnaires.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Authors' contributions

All authors declare that they have read and approved the final version of this article. The total percentage of contribution for the conceptualization, preparation and correction of this article was equitable (14.286 % each).

Statement on data availability

Data supporting the results of this study will be made available by the corresponding author (Z.A.A), upon reasonable request.

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

References

Bachillerato y Licenciantura en la Enseñanza de la Matemática (BLEM). (2017). *Plan de estudios de Bachillerato y Licenciatura en Enseñanza de la Matemática*. Heredia, Costa Rica: Universidad Nacional. http://www.matematica.una.ac.cr/

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Berkowitz, R. (2021). School climate and the socioeconomic literacy achievement gap: Multilevel analysis of compensation, mediation, and moderation models. *Children and Youth Services Review,* 130, 1-10. https://doi.org/10.1016/j.childyouth.2021.106238

Berkowitz, R., Glickman, H., Benbenishty, R., Ben-Artzi, E., Raz, T., Lipshtat, N. & Astor, R. A. (2015). Compensating, Mediating, and Moderating Effects of School Climate on Academic Achievement Gaps in Israel. *Teachers College Record*, *117*(7), 1-34. https://doi.org/10.1177/016146811511700703

Bong, M. (2004). Academic Motivation in Self-Efficacy, Task Value, Achievement Goal Orientations, and Attributional Beliefs. *The Journal of Educational Research*, *97*(6), 287-298. https://doi.org/10.3200/JOER.97.6.287-298

Chamizo-Nieto, M.T., Arrivillaga, C., Rey, L. & Extremera, N. (2021). The Role of Emotional Intelligence, the Teacher-Student Relationship, and Flourishing on Academic Performance in Adolescents: A Moderated Mediation Study. *Front. Psychol, 12.* https://doi.org/10.3389/fpsyg.2021.695067

Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological bulletin*, *136*(1), 103-127. http://dx.doi.org/10.1037/a0018053

Expósito, F., Moya, M. & Glick, P. (1998). Sexismo ambivalente: medición y correlatos. *Revista de Psicología Social*, *13,* 159-169. https://doi.org/10.1174/021347498760350641

Fennema, E. & Sherman, J. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes toward the Learning of Mathematics by Females and Males. Journal for Research in Mathematics Education, 7(5), 324-326. https://doi.org/10.5951/jresematheduc.7.5.0324

Finch, W. H., Bolin, J. E. & Kelley, K. (2014). Multilevel modeling using R. Crc Press. https://doi.org/10.1201/b17096

Forgasz, H. J., Leder, G. C. & Gardner, P. L. (1999). The Fennema-Sherman Mathematics as a Male Domain Scale Reexamined. *Journal for Research in Mathematics Education, 30*(3), 342-348. https://doi.org/10.2307/749839

Glick, P. & Fiske, S. T. (1996). The Ambivalent Sexism Invntory: Differentiating hostile and benevolent sexism. Journal of Personality and Social Psychology. 70. 491-512. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P. & Fiske, S. T. (2011). Ambivalent sexism revisited. Psychology of Women Quarterly, 35(3), 530-535. https://doi.org/10.1177/0361684311414832

Guiso, L., Monte, F., Sapienza, P. & Zingales, L. (2008). Diversity. Culture, gender, and math. *Science*, *320*(5880), 1164-1165. https://doi.org/10.1126/science.1154094

Huang, C. (2013). Gender differences in academic self-efficacy: a meta-analysis. *European Journal of Psychology of Education, 28*, 1-35. https://doi.org/10.1007/s10212-011-0097-y

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B. & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494-495. https://doi.org/10.1126/science.1160364

Jacobs, J. E. & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, *63*, 932-944. https://doi.org/10.1037/0022-3514.63.6.932

Jordan, K. & Carden, R. (2017). Self-efficacy and gender in STEM majors. *Modern Psychological Studies, 22*(2), 60-64. https://scholar.utc.edu/mps/vol22/iss2/8

Kaplan, D. (2014). *Bayesian Statistics for the Social Sciences*. New York, New York: Guilford Press.

Leaper, C. & Brown, C. S. (2014). Chapter six - Sexism in Schools. In L. S. Liben & R. S. B. T.-A. in C. D. and B. Bigler (Eds.), *The Role of Gender in Educational Contexts and Outcomes*, *47,* 189-223. https://doi.org/10.1016/bs.acdb.2014.04.001

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

14

Lindberg, S. M., Hyde, J. S., Petersen, J. L. & Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychological bulletin*, *136*(6), 1123-1135. https://doi.org/10.1037/a0021276

Lynn, R. & Mikk, J. (2009). National IQs predict educational attainment in math, reading and science across 56 nations. *Intelligence*, *37*(3), 305-310. https://doi.org/10.1016/j.intell.2009.01.002

Martínez, F., Lizasoain, L., Castro, M. & Joaristi, L. M. (2017). Selección de escuelas de alta y baja eficacia en Baja California (México). *Revista Electrónica de Investigación Educativa*, *19*(2), 38-53. https://doi.org/10.24320/redie.2017.19.2.960

Mella, N., Pansu, P., Batruch, A., Bressan, M., Bressoux, P., Brown, G., Butera, F., Cherbonnier, A., Darnon, C.,Demolliens, M., De Place, A.L., Huguet, P., Jamet, E., Martinez, R., Mazenod, V., Michinov, E., Michinov, N., Poletti, C., Régner, I., Riant, M., Robert, A., Rudmann, O., Sanrey, C., Stanczak, A.,Visintin, E.P., Vives, E. & Desrichard, O. (2021). Socio-Emotional Competencies and School Performance in Adolescence: What Role for School Adjustment? *Front. Psycholy*. https://doi.org/10.3389/fpsyg.2021.640661

Mena, P. (2015). Desarrollo en la prueba nacional de bachillerato de Matemática: una necesidad. *Cuadernos de Investigación y Formación en Educación Matemática*, *10* (13), 53-66. http://funes.uniandes.edu.co/21494/1/Mena-2015Desarrollo.pdf

Montero, E., Castelain, T., Moreira, T., Alfaro, L., Cerdas, D., García A. & Roldán, G. (2013). Evidencias iniciales de validez de criterio de los resultados de una Prueba de razonamiento con figuras para la selección de estudiantes indígenas para la Universidad de Costa Rica y el Instituto Tecnológico de Costa Rica. *Revista Educación, 37*(2), 103-117. https://doi.org/10.15517/revedu.v37i2.12928

Montero, E., Rojas, G., Negrín, M. & Francis, S. (2015). Efecto de una capacitación sobre los puntajes de la prueba de admisión de la Universidad de Costa Rica: una aproximación bayesiana. *Actualidades en Psicología*, *29*(119). San José, Costa Rica: Universidad de Costa Rica. http://dx.doi.org/10.15517/ap.v29i119.19283

Montero-Rojas, E., Moreira-Mora, T., Zamora-Araya, J.A. & Smith-Castro, V. (2021). Una nueva mirada teórica y metodológica a las diferencias de género en pruebas de matemática: razonamiento, actitudes psicosociales y modelos multinivel. *Revista Educare, 25*(1), 1-21. http://dx.doi.org/10.15359/ree.25-1.8

Moreira-Mora, T., Zamora-Araya, J. A., Smith-Castro, V. & Montero-Rojas, E. (2021). Rezago en asignaturas de contenido matemático en estudiantes de último año: Analizando resultados de colegios en Territorios Indígenas y del Gran Área Metropolitana de Costa Rica. *Revista Educación,45*(1). https://doi.org/10.15517/revedu.v45i1.41317

Murillo, F. J. & Román, M. (2009). Mejorar el desempeño de los estudiantes de América Latina: algunas reflexiones a partir de los resultados del SERCE. *Revista Mexicana de Investigación Educativa, 14*(41), 451-484. http://www.scielo.org.mx/pdf/rmie/v14n41/v14n41a6.pdf

Navarro, G., Flores-Oyarzo, G. & González, M.G. (2019). Diferencias por sexo en el nivel de Autoeficacia percibida en una muestra de estudiantes de la Provincia de Concepción ¿Qué papel juegan los roles de género en la educación? *Revista de Educación Inclusiva*, *12*(1), 205-224. https://revistaeducacioninclusiva.es/index.php/REI/article/view/389/374

OCDE. (2014). *Panorama de la educación 2013: Indicadores de la OCDE*. Santillana, España, Madrid. https://doi.org/10.1787/eag-2013-es

OCDE. (2015). *Panorama de la educación 2014: Indicadores de la OCDE*. Fundación Santillana, Madrid. https://doi.org/10.1787/eag-2014-es

O'Malley, M., Voight, A., Renshaw, T. L., & Eklund, K. (2015). School climate, family structure, and academic achievement: a study of moderation effects. *School Psychology Quarterly, 30*(1), 142. https://doi.org/10.1037/spq0000076

Recber, S., Isiksal, M. & Koç, Y. (2018). Investigating self-efficacy, anxiety, attitudes and mathematics achievement regarding gender and school type. *Anales de Psicología*, *34*(1), 41-51. https://dx.doi.org/10.6018/analesps.34.1.229571

Ruiz, L. D., McMahon, S. D. & Jason, L. A. (2018). El papel del contexto del vecindario y el clima escolar en el rendimiento académico a nivel escolar. *Revista estadounidense de psicología comunitaria* , *61* (3-4), 296-309. https://doi.org/10.1002/ajcp.12234

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

15

Ruiz-Benito, P., Andivia, E., Archambeaou, J., Astigarraga, J., Barrientos, R., Cruz-Alonso, V., Florencio, M., Gómez, D., Martínez-Baroja, Zoë, P.Q., Santos, A., Velado, E., Villén-Pérez, S. & Morales-Castilla, I. (2018). Ventajas de la estadística bayesiana frente a la frecuentista: ¿por qué nos resistimos a usarla? *Ecosistemas*, *27*(2), 136-139. https://hal.inrae.fr/hal-02617924/document

Schwarzer, R., Bäßler, J., Kwiatek, P., Schröder, K. & Zhang, J. X. (1997). The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the General Self-efficacy Scale. *Applied Psychology: An International Review*, *46*(1), 69-88. https://doi.org/10.1111/j.1464-0597.1997.tb01096.x

Smith-Castro, V., Montero-Rojas, E., Moreira-Mora, T. & Zamora-Araya, J. A. (2019). Expected and unexpected effects of sexism on women´s mathematics performance, *Interamerican Journal of Psychology, 53*(1), 28-44. https://doi.org/10.30849/rip/ijp.v53i1.905

Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4-28. https://doi.org/10.1006/jesp.1998.1373

Steele, C. M., Spencer, S. J. & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), Advances in experimental social psychology, 34, 379-440. https://doi.org/10.1016/S0065-2601(02)80009-0

UNESCO. (2020). *¿Qué se espera que aprendan los estudiantes de América Latina y el Caribe? Análisis curricular del Estudio Regional Comparativo y Explicativo (ERCE 2019).* UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean. https://unesdoc.unesco.org/ark:/48223/pf0000373982

Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174-1204. https://doi.org/10.1037/a0036620

Walton G. M. & Spencer S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, 20. https://doi.org/10.1111/j.1467-9280.2009.02417.x

Zamora. J. A. (2020). Las actitudes hacia la matemática, el desarrollo social, el nivel educativo de la madre y la autoeficacia como factores asociados al rendimiento académico en la matemática. Uniciencia, *34*(1), 74-87. https://doi.org/10.15359/ru.34-1.5

Zell, E., Strickhouser, J. E., Lane, T. N. & Teeter, S. R. (2016). Mars, Venus, or Earth? Sexism and the exaggeration of psychological gender differences. *Sex Roles*, 75, 287-300. https://link.springer.com/article/10.1007/s11199-016-0622-1

UNICIENCIA Vol. 36, N° 1, pp. 1-17. January-December, 2022 • ⊕ www.revistas.una.ac.cr/uniciencia • ✉ revistauniciencia@una.cr

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

16

Appendix 1
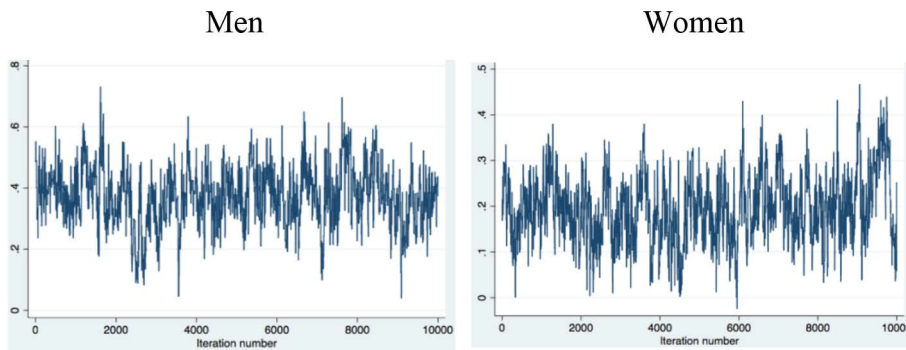
Men   Women

Figure 1. *Trace according to sex for the estimated model. 2021*
Source: Calculated by the authors during this investigation.

Figure 1 shows the traces of analyses for men and women, in which stationary behavior of the chains is observed with 10,000 iterations after burning the first 2,500 values, showing a convergence in model estimates.

José Andrey Zamora-Araya • Eiliana Montero-Rojas • Vanessa Smith-Castro • Tania Elena Moreira-Mora • Pablo Zamora-Calvo • Karla Quintero-Arias • Stefani Matarrita-Muñoz

17