

# Scientia et PRAXIS

*Vol. 06 No. 11-2026*

*eISSN 2954-4041*



**AMIDI**  
Academia Mexicana  
de Investigación y Docencia  
en Innovación



**Volumen 06, Número 11**

**Enero-Junio 2026**

**eISSN: 2954-4041**

<https://doi.org/10.55965/setp.6.11>

**-Director AMIDI-**

**Dr. Juan Mejía-Trejo**

Universidad de Guadalajara, Guadalajara, Jalisco, México.

**Miembros del Consejo Editorial:**

**-Editor en Jefe-**

**Dr. Carlos Gabriel Borbón-Morales**

Centro de investigación en Alimentación y Desarrollo (CIAD-SECIHTI) Hermosillo, Sonora

**-Editor Asociado-**

**Dr. Carlos Omar Aguilar-Navarro.**

Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco A.C. (CIATEJ-SECIHTI), Guadalajara, Jalisco, México.

**-Comité Científico-**

**Dr. Guillermo Peinado**

Universidad Nacional de Rosario, Rosario, Santa Fé, Argentina.

**Dr. Claudia De-Fuentes.**

Saint Mary's University, Halifax, Nova Scotia, Canadá.

**Dr. Jaime Antero Arango-Marin.**

Universidad Católica Luis Amigó, Medellín, Antioquia, Colombia.

**Dr. Abu Waheeduzzaman**

Texas. A&M University-Corpus Christi, EUA.

**Dr. Héctor Ortiz-Cano**

Northwestern University. Illinois, EUA.

**Dr. Ángel Rodríguez-Bravo.**

Asociación Científica para la Evaluación y Medición de los Valores Humanos (AEVA), Barcelona, España.

**Dra. Norminanda Montoya-Vilar. ORCID.**

Universidad Politécnica de Cartagena, España.

**Dra. Antonia Madrid-Guijarro.**

**Dr. Domingo García-Pérez de Lema.**

Universidad Politécnica de Cartagena, España.

**Dra. Laura Nieves -Sierra García.**

Universidad Pablo de Olavide, Sevilla, España.

**Dr. Jiachen Hou**

University of Bradford. Reino Unido.

**Dr. Yari Borbón-Gálvez**

Universita Carlo Cattaneo: Castellanza, Lombardia, Italia.

**Dr. Miguel Ángel Martínez-Télez**

**Dr. José Angel Vega-Noriega**

Centro de Investigación en Alimentación y Desarrollo (CIAD- SECIHTI), Hermosillo, Sonora, México.

**Dr. Ramón Jaime Holguín-Peña**

Centro de Investigaciones Biológicas del Noroeste S.C. (CIBNOR- SECIHTI).La Paz, Baja California Sur, México.

**Dr. Héctor González-Ocampo.**

Centro Interdisciplinario de Investigación para el Desarrollo integral Regional (CIIDIR). Instituto Politécnico Nacional (IPN).Guasave, Sinaloa, México.

**Dr. Eduardo Morales-Sánchez.**

Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada (CICATA).Querétaro, Querétaro, México.

**Dra. Emma Regina Morales García de Alba.**

Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO). Guadalajara, Jalisco, México.

**Dr. Gerardo Rodríguez-Barba**

Centro de Investigación y Asistencia Técnica del Estado de Querétaro especializado en Manufactura Avanzada y Procesos Industriales (CIATEQ-SECIHTI). Guadalajara, Jalisco, México.

**Dr. Enrique Saldívar-Guerra**

Centro de Investigación de Química Aplicada (CIQA- SECIHTI), Saltillo, Coahuila, México.

**Dra. Paulina Elisa Lagunes-Navarro**

Centro de Investigación e Innovación en TIC (INFOTEC- SECIHTI), Ciudad de México, México.

**Dra. América Berenice Morales-Díaz**

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Unidad (CINVESTAV) Saltillo, Coahuila, México.

**Dr. Clara Galindo-Sánchez**

Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE- SECIHTI), Baja California, México.

**Dr. Antonio Aguilera-Ontiveros**

Colegio de San Luis (COLSAN- SECIHTI), San Luis de Potosi, San Luis Potosi, México.

**Dr. Yanga Villagomez-Velázquez**

Colegio de Michoacán (COLMICH- SECIHTI), Zamora, Michoacán,. México.

**Dr. Héctor Medina-Miranda**

Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS- SECIHTI), Guadalajara, Jalisco, México

**Dr. José Tuxpan-Vargas**

Instituto Potosino de Investigación Científica y Tecnológica (IPICYT-SECIHTI). San Luis Potosi, San Luis Potosí, México.

**Dr. Miguel Eduardo Equihua-Zamora**

Instituto de Ecología (INECOL-SECIHTI) Xalapa, Veracruz, México.

**Dr. Luis Sáenz-Carbonell**

Centro de Investigación Científica de Yucatán (CICY- SECIHTI) Merida, Yucatan, México.

**Dr. Alejandro Morón-Ríos**

El Colegio de la Frontera Sur (ECOSUR- SECIHTI), Campeche, Campeche, México.

**Dr. Jorge Castañeda Zavala**

Instituto de Investigaciones Dr. José María Luis Mora (SECIHTI) Ciudad de México, México.

**Dra. María del Rosio Barajas-Escamilla**

El Colegio de la Frontera Norte (COLEF- SECIHTI), Tijuana, Baja California, México.

**Dra. Helena Cotler**

CentroGeo (SECIHTI), Ciudad de México, México.

**Dra. Elia Marum Espinosa (CUCEA-Cátedra UNESCO)**

**Dr. César Omar Mora-Pérez (CUCEA)**

**Dr. Antonio Ruiz-Porras (CUCEA)**

**Dr. Jaime Antonio Preciado-Coronado (CUCSH)**

Universidad de Guadalajara (UdeG), Guadalajara, Jalisco, México.

## Carta Editorial

Volumen 06, Número 11 | Enero–Junio de 2026

La Academia Mexicana de Investigación y Docencia en Innovación (AMIDI), a través del Consejo Editorial de la revista *Scientia et PRAXIS*, presenta el Volumen 06, Número 11, correspondiente al período **enero-junio de 2026**, como edición regular. Esta entrega reúne trabajos científicos originales e inéditos que exploran cómo la actividad multidisciplinaria se constituye como motor de la innovación con impacto en el desarrollo sostenible y la transformación social. Las contribuciones incluidas en este número destacan por su articulación entre el conocimiento teórico (*Scientia*) y su aplicación práctica (*Praxis*), en consonancia con los Objetivos de Desarrollo Sostenible (ODS).

Los artículos que conforman esta edición son los siguientes:

**1 Hacia una innovación de proceso de evaluación docente sostenible: revisión sistemática de aplicaciones de IA, Ciencia de Datos y PLN en educación superior.**

*(Artículo escrito en español).*

**Autores:** Cristian Ulises Barenca-Sotelo, Ma. Del Rocío Maciel-Arellano, Víctor Manuel Larios-Rosillo.

**Universidad de Guadalajara, Guadalajara, Jalisco.**

**Reseña:** El artículo analiza, mediante revisión sistemática 2019–2024, el uso de inteligencia artificial, ciencia de datos y procesamiento de lenguaje natural en la evaluación docente universitaria. Su principal aporte es desplazar la evaluación tradicional basada en promedios y encuestas cerradas hacia modelos analíticos capaces de interpretar comentarios estudiantiles, identificar patrones de desempeño y apoyar decisiones institucionales. La propuesta se vincula con innovación de proceso, Manual de Oslo y ODS 4 y 9, destacando pertinencia para universidades públicas multicampus. Su valor reside en integrar tecnología, sostenibilidad y mejora académica, aunque requiere validación empírica posterior en contextos institucionales específicos.

**DOI:** <https://doi.org/10.55965/setp.6.11.a1>

## **2 Emociones y Machine Learning en la innovación del mercado sostenible de vehículos seminuevos en México.**

*(Artículo escrito en español).*

**Autor: Francisco Jacobo Murillo-López.**

**Universidad Autónoma de Aguascalientes, Aguascalientes, México.**

**Reseña:** *El artículo analiza cómo las emociones del consumidor influyen en la lealtad dentro del mercado mexicano de vehículos seminuevos, integrando neuroeconomía, comportamiento del consumidor y machine learning. Con una muestra de 1,000 compradores en Aguascalientes, compara regresión logística y Random Forest, mostrando mayor capacidad predictiva del modelo no lineal. Identifica satisfacción y seguridad como impulsores de lealtad, mientras miedo y confusión actúan como barreras. Su aporte abona principalmente al ODS 12, por consumo responsable y economía circular, y al ODS 9, por innovación de proceso basada en analítica predictiva para decisiones comerciales sostenibles en mercados automotrices regionales contemporáneos.*

**DOI: <https://doi.org/10.55965/setp.6.11.a2>**

## **3 Género, cultura de paz y participación ciudadana como innovación social basada en evidencia para la gobernanza sostenible.**

*(Artículo escrito en inglés)*

**Autores: Tania Marcela Hernández-Rodríguez, César Omar Mora-Pérez**

**Universidad de Guadalajara, Guadalajara, Jalisco, México.**

**Reseña:** *El artículo analiza cómo el género modera la relación entre cultura de paz, participación ciudadana y percepción de conflictos vecinales en comunidades urbanas del Área Metropolitana de Guadalajara. Con enfoque cuantitativo transversal y muestra de 229 residentes, utiliza escalas validadas, análisis factorial exploratorio, alfa de Cronbach, correlaciones de Pearson y regresión con interacción. Su aporte consiste en fundamentar intervenciones de innovación social para gobernanza urbana sostenible con perspectiva de género, alineadas con los ODS 5, 11 y 16. Destaca evidencia empírica relevante, aunque su alcance se limita por diseño transversal y concentración geográfica metropolitana.*

**DOI: <https://doi.org/10.55965/setp.6.11.a3>**

## **4 CAITIZEN extendido: estudio PLS-SEM de la innovación en ciudadanía sostenible asistida por IA**

*(Artículo escrito en inglés)*

**Autor: Juan Mejía-Trejo**

**Universidad de Guadalajara, Guadalajara, Jalisco, México.**

**Reseña:** *El artículo valida cuantitativamente el modelo extendido CAITIZEN mediante PLS-SEM, proponiendo una estructura explicativo-predictiva para la ciudadanía sostenible asistida por inteligencia artificial. Su aporte central es convertir un marco cualitativo previo en un modelo medible, donde la alfabetización crítica en IA predice conciencia ética, justicia de datos, colaboración humano-IA y transparencia metacognitiva. Los resultados confirman efectos relevantes hacia CAITIZEN, salvo el efecto directo de responsabilidad ética. La investigación fortalece la innovación educativa alineada con los ODS 4 y 9, aunque reconoce limitaciones por muestreo no probabilístico, diseño transversal y muestra estudiantil, requiriendo validaciones futuras comparativas en diversos contextos institucionales universitarios.*

**DOI:** <https://doi.org/10.55965/setp.6.11.a4>

Estos artículos reflejan una **contribución significativa al desarrollo sostenible** desde perspectivas innovadoras y multidisciplinarias. Los **autores**, las **autoridades de AMIDI** y el **Consejo Editorial** de la revista científica *Scientia et PRAXIS* que participaron en esta obra desean que los lectores encuentren información **accesible, rigurosa y útil para sus propósitos formativos, investigativos o profesionales**. Asimismo, los invitan cordialmente a postular sus propios trabajos para futuras ediciones como esta, en las que se analizan críticamente los problemas de nuestro país y del mundo, y se proponen soluciones fundamentadas desde una perspectiva académica comprometida con la transformación social.

**Dr. Juan Mejía-Trejo**  
Director  
Academia Mexicana de  
Investigación y Docencia en  
Innovación (AMIDI)  
Junio 2026, Zapopan, Jalisco,  
México

**Dr. Carlos G. Borbón-Morales**  
Editor en Jefe  
*Scientia et PRAXIS*  
Academia Mexicana de Investigación  
y Docencia en Innovación (AMIDI)  
Junio 2026, Zapopan, Jalisco, México

## Editorial Letter

Volume 06, Number 11 | January–June 2026

The **Academia Mexicana de Investigación y Docencia en Innovación (AMIDI)**, through the Editorial Board of the journal *Scientia et PRAXIS*, presents **Volume 06, Number 11**, corresponding to the **January–June 2026** period, as a regular issue. This edition features original and unpublished scientific works that explore how multidisciplinary activity serves as a driving force for innovation with impact on sustainable development and social transformation. The contributions included in this volume stand out for their articulation between theoretical knowledge (*Scientia*) and practical application (*Praxis*), in alignment with the **Sustainable Development Goals (SDGs)**.

The articles presented in this issue are as follows:

**1 Toward a Sustainable Innovation in the Faculty Evaluation Process: A Systematic Review of AI, Data Science, and NLP Applications in Higher Education.**

*(Article written in Spanish)*

**Authors: Cristian Ulises Barenca-Sotelo, Ma. Del Rocío Maciel-Arellano, Víctor Manuel Larios-Rosillo**

**Universidad de Guadalajara, Guadalajara, Jalisco.**

**Summary:** *The article analyzes, through a 2019–2024 systematic review, the use of artificial intelligence, data science, and natural language processing in university teaching evaluation. Its main contribution lies in shifting traditional evaluation based on averages and closed-ended surveys toward analytical models capable of interpreting student comments, identifying performance patterns, and supporting institutional decision-making. The proposal is linked to process innovation, the Oslo Manual, and SDGs 4 and 9, highlighting its relevance for multicampus public universities. Its value resides in integrating technology, sustainability, and academic improvement, although further empirical validation is required in specific institutional contexts.*

**DOI: <https://doi.org/10.55965/setp.6.11.a1>**

## 2 Emotions and Machine Learning in Innovation within Mexico's Sustainable Used-Vehicle Market.

*(Article written in Spanish)*

**Authors: Francisco Jacobo Murillo-López.**

**Universidad Autónoma de Aguascalientes, Aguascalientes, México.**

**Summary:** *The article analyzes how consumer emotions influence loyalty within Mexico's used-vehicle market, integrating neuroeconomics, consumer behavior, and machine learning. Based on a sample of 1,000 buyers in Aguascalientes, it compares logistic regression and Random Forest, showing the greater predictive capacity of the nonlinear model. It identifies satisfaction and safety as drivers of loyalty, while fear and confusion operate as barriers. Its contribution is mainly aligned with SDG 12, through responsible consumption and circular economy, and with SDG 9, through process innovation based on predictive analytics for sustainable business decision-making in contemporary regional automotive markets.*

**DOI: <https://doi.org/10.55965/setp.6.11.a2>**

## 3 Gender, Culture of Peace, and Citizen Participation as Evidence-Based Social Innovation for Sustainable Governance.

*(Article written in English).*

**Authors: Tania Marcela Hernández-Rodríguez, César Omar Mora-Pérez**

**Universidad de Guadalajara, Guadalajara, Jalisco, México.**

**Summary:** *The article analyzes how gender moderates the relationship between culture of peace, citizen participation, and the perception of neighborhood conflicts in urban communities within the Guadalajara Metropolitan Area. Using a cross-sectional quantitative approach and a sample of 229 residents, it applies validated scales, exploratory factor analysis, Cronbach's alpha, Pearson correlations, and interaction regression. Its contribution lies in supporting evidence-based social innovation interventions for sustainable urban governance with a gender perspective, aligned with SDGs 5, 11, and 16. The study provides relevant empirical evidence, although its scope is limited by its cross-sectional design and metropolitan geographic concentration.*

**DOI: <https://doi.org/10.55965/setp.6.11.a3>**

## 4 Extended CAITIZEN: A PLS-SEM Study of AI-Assisted Sustainable Citizenship Innovation.

*(Article written in English).*

**Author: Juan Mejía-Trejo**

**Universidad de Guadalajara, Guadalajara, Jalisco, México.**

***Summary:** The article quantitatively validates the extended CAITIZEN model through PLS-SEM, proposing an explanatory-predictive structure for AI-assisted sustainable citizenship. Its central contribution is the transformation of a previous qualitative framework into a measurable model, in which critical AI literacy predicts ethical awareness, data justice, human-AI collaboration, and metacognitive transparency. The results confirm relevant effects on CAITIZEN, except for the direct effect of ethical responsibility. The study strengthens educational innovation aligned with SDGs 4 and 9, while acknowledging limitations related to non-probabilistic sampling, cross-sectional design, and a student-based sample, thus requiring future comparative validations across diverse university institutional contexts.*

**DOI: <https://doi.org/10.55965/setp.6.11.a4>**

These articles reflect a **significant contribution to sustainable development** from **innovative and multidisciplinary perspectives**. The **authors**, the **AMIDI authorities**, and the **Editorial Board** of *Scientia et PRAXIS* who contributed to this volume sincerely hope that readers will find the information **accessible, rigorous, and useful** for their educational, research, or professional objectives. They also warmly **invite scholars and professionals** to submit their own work to future issues like this one—dedicated to critically examining national and global challenges and proposing evidence-based, socially committed solutions from an academic perspective.

**Dr. Juan Mejía-Trejo**  
Director  
Academia Mexicana de  
Investigación y Docencia en  
Innovación (**AMIDI**)  
June 2026, Zapopan, Jalisco, México

**Dr. Carlos G. Borbón-Morales**  
Editor-in-Chief  
***Scientia et PRAXIS***  
Academia Mexicana de Investigación y  
Docencia en Innovación (**AMIDI**)  
June 2026, Zapopan, Jalisco, México

## Contenido *Content*

- 1 Hacia una innovación de proceso de evaluación docente sostenible: Revisión Sistemática de Aplicaciones de IA, Ciencia de Datos y PLN en Educación Superior.**

*Toward a Sustainable Innovation in the Faculty Evaluation Process: A Systematic Review of AI, Data Science, and NLP Applications in Higher Education*

**Cristian Ulises Barenca-Sotelo**

**Ma. Del Rocío Maciel-Arellano**

**Víctor Manuel Larios-Rosillo**

Universidad de Guadalajara, Guadalajara, Jalisco.

- 25 Emociones y Machine Learning en la innovación del mercado sostenible de vehículos seminuevos en México.**

*Emotions and Machine Learning in Innovation within Mexico's Sustainable Used-Vehicle Market.*

**Francisco Jacobo Murillo-López.**

Universidad Autónoma de Aguascalientes, Aguascalientes, México.

- 59 Gender, Culture of Peace, and Citizen Participation as Evidence-Based Social Innovation for Sustainable Governance.**

*Género, cultura de paz y participación ciudadana como innovación social basada en evidencia para la gobernanza sostenible.*

**Tania Marcela Hernández-Rodríguez, César Omar Mora-Pérez**

Universidad de Guadalajara, Guadalajara, Jalisco, México.

- 86 Extended CAITIZEN: A PLS-SEM Study of AI-Assisted Sustainable Citizenship Innovation**

*CAITIZEN extendido: estudio PLS-SEM de la innovación en ciudadanía sostenible asistida por IA*

**Juan Mejía-Trejo**

Universidad de Guadalajara, Guadalajara, Jalisco, México.

# Scientia et PRAXIS

Vol. 06. No.11. Ene-Jun (2026): 1-24

<https://doi.org/10.55965/setp.6.11.a1>

eISSN: 2954-4041

## Hacia una innovación de proceso de evaluación docente sostenible: Revisión Sistemática de Aplicaciones de IA, Ciencia de Datos y PLN en Educación Superior

## Toward a Sustainable Innovation in the Faculty Evaluation Process: A Systematic Review of AI, Data Science, and NLP Applications in Higher Education

**Cristian Ulises Barenca-Sotelo.** ORCID [0009-0000-5091-1900](https://orcid.org/0009-0000-5091-1900)

Universidad de Guadalajara, Centro Universitario de Ciencias Económico Administrativas,  
Zapopan, Jalisco, México

e-mail: [cristian.barenca1897@alumnos.udg.mx](mailto:cristian.barenca1897@alumnos.udg.mx)

**Ma. Del Rocío Maciel-Arellano.** ORCID [0000-0002-5548-2073](https://orcid.org/0000-0002-5548-2073)

Universidad de Guadalajara, Centro Universitario de Ciencias Económico Administrativas,  
Zapopan, Jalisco, México

e-mail: [ma.maciel@academicos.udg.mx](mailto:ma.maciel@academicos.udg.mx)

**Víctor Manuel Larios-Rosillo.** ORCID [0000-0002-2899-724X](https://orcid.org/0000-0002-2899-724X)

Universidad de Guadalajara, Centro Universitario de Ciencias Económico Administrativas,  
Zapopan, Jalisco, México

e-mail: [victor.larios@academicos.udg.mx](mailto:victor.larios@academicos.udg.mx)

**Palabras Clave:** evaluación docente, educación superior, inteligencia artificial, procesamiento de lenguaje natural, revisión sistemática.

**Keywords:** faculty evaluation, higher education, artificial intelligence, natural language processing, systematic review.

---

**Recibido:** 12-Nov-2025; **Aceptado:** 22-Feb-2026

## **RESUMEN**

**Contexto.** La evaluación docente en universidades públicas multicampus depende aún de promedios y encuestas cerradas, limitadas para captar la complejidad del desempeño académico. Ante ello, este estudio presenta una revisión sistemática (2019–2024) sobre el uso de la Inteligencia Artificial (IA) y el Procesamiento de Lenguaje Natural (PLN) en procesos de evaluación docente universitaria, siguiendo como innovaciones de proceso enunciadas en el Manual de Oslo, alineadas con los Objetivos de Desarrollo Sostenibles (ODS) 4 y 9.

**Problema.** Los sistemas tradicionales en instituciones de alta matrícula presentan baja precisión analítica, escaso aprovechamiento de datos cualitativos y retroalimentación tardía. Ante esto, se plantea la pregunta de investigación: ¿qué técnicas y algoritmos reporta la literatura para el análisis integrado de datos cuantitativos y cualitativos en la evaluación docente universitaria?

**Objetivo.** Analizar críticamente el uso de la ciencia de datos y la IA —especialmente el PLN— para mejorar la retroalimentación docente, superando las limitaciones de los enfoques tradicionales y promover procesos evaluativos oportunos, profundos y escalables.

**Metodología.** Se siguió el protocolo PRISMA mediante búsquedas con operadores booleanos en Scopus, Web of Science, IEEE Xplore y ERIC. Tras aplicar criterios de inclusión y un proceso de revisión por pares en dos fases, se analizaron 17 estudios publicados entre 2019 y 2024.

**Hallazgos.** Las técnicas como el análisis de sentimientos, el modelado de tópicos (LDA) y modelos de lenguaje de gran escala (LLM) —notablemente DistilBERT, con precisiones cercanas al 93%— superan consistentemente a los métodos tradicionales en la gestión de grandes volúmenes de información.

**Originalidad.** Radica en la integración de literatura dispersa sobre IA y PLN en educación superior bajo un marco de innovación de procesos, articulando rigor técnico, empírico y ético.

**Conclusiones y Limitaciones.** Se concluye que los modelos avanzados de IA y PLN poseen un alto potencial para transformar la evaluación en redes universitarias latinoamericanas al permitir retroalimentación personalizada. No obstante, persisten desafíos en la interpretabilidad de modelos, mitigación de sesgos, escasez de datos etiquetados en español y resistencia institucional, lo que abre relevantes líneas para futuras investigaciones y desarrollos aplicados.

## ABSTRACT

**Context.** Teaching evaluation in public multicampus universities still relies on averages and closed-ended surveys, which are limited in capturing the complexity of academic performance. In response, this study presents a systematic review (2019–2024) on the use of Artificial Intelligence (AI) and Natural Language Processing (NLP) in university teaching evaluation processes, conceptualized as process innovations from Oslo Manual and aligned with Sustainable Development Goals (SDGs) 4 and 9.

**Problem.** Traditional systems in high-enrollment institutions exhibit low analytical precision, limited use of qualitative data, and delayed feedback. Accordingly, the following research question is posed: which techniques and algorithms does the literature report for the integrated analysis of quantitative and qualitative data in university teaching evaluation?

**Purpose.** To critically analyze the use of data science and AI—particularly NLP—to enhance teaching feedback, overcoming the limitations of traditional approaches and promoting timely, in-depth, and scalable evaluation processes.

**Methodology.** The PRISMA protocol was followed through searches using Boolean operators in Scopus, Web of Science, IEEE Xplore, and ERIC. After applying inclusion criteria and a two-phase peer-review process, 17 studies published between 2019 and 2024 were analyzed.

**Findings.** Techniques such as sentiment analysis, topic modeling (LDA), and large language models (LLMs)—notably DistilBERT, with accuracy levels close to 93%—consistently outperform traditional methods in managing large volumes of information.

**Originality.** The study’s originality lies in integrating dispersed literature on AI and NLP in higher education within a coherent process-innovation framework, combining technical, empirical, and ethical rigor.

**Conclusions and Limitations.** Advanced AI and NLP models show high potential to transform evaluation in Latin American university networks by enabling personalized feedback. However, challenges remain regarding model interpretability, bias mitigation, the scarcity of labeled data in Spanish, and institutional resistance to change, opening relevant avenues for future research and applied developments.

## **1. INTRODUCCIÓN**

En las instituciones de educación superior de América Latina, particularmente en universidades públicas de gran escala organizadas en redes universitarias y con estructuras multicampus, la evaluación del desempeño docente constituye un proceso estratégico para el aseguramiento de la calidad educativa. No obstante, dichos procesos continúan apoyándose mayoritariamente en modelos tradicionales basados en promedios numéricos y cuestionarios cerrados, los cuales resultan insuficientes para capturar la complejidad de la práctica docente y para aprovechar de manera sistemática la información cualitativa generada por el estudiantado. En este contexto, universidades con altos volúmenes de matrícula y una amplia diversidad disciplinaria —como aquellas que conforman la Red Universitaria de la Universidad de Guadalajara (**UDG**)— enfrentan desafíos significativos asociados al procesamiento masivo de datos, la oportunidad de la retroalimentación y la toma de decisiones informadas.

Centros universitarios como el de Ciencias Económico Administrativas (**CUCEA**) ilustran estas tensiones, al operar sistemas de evaluación que deben atender millones de registros por ciclo académico, situación que refleja problemáticas comunes a numerosas instituciones latinoamericanas y pone de manifiesto la necesidad de transitar hacia modelos de evaluación docente apoyados en innovaciones de proceso basadas en ciencia de datos, Inteligencia Artificial y Procesamiento de Lenguaje Natural.

La evaluación docente en el ámbito universitario ha dependido históricamente de métodos tradicionales que priorizan promedios simples y encuestas de preguntas cerradas, si bien estos enfoques permiten generar métricas básicas, resultan insuficientes para capturar la complejidad del desempeño académico y ofrecer retroalimentación significativa y oportuna.

Asimismo, en la **UDG**, y particularmente en el **CUCEA**, estas limitaciones se evidencian en el procesamiento de más de 2'354,631 registros por ciclo, lo que retrasa la toma de decisiones pedagógicas y conduce a análisis superficiales que no reflejan el desempeño real del profesorado.

Ahora bien, frente a estas deficiencias, la **IA**, Ciencia de Datos y el **PLN** emergen como enfoques innovadores de procesos organizacionales, según el Manual de Oslo, ya que estas tecnologías permiten automatizar el análisis de grandes volúmenes de datos y transformar comentarios cualitativos de estudiantes en información útil mediante técnicas como análisis de sentimientos y modelos de lenguaje (**LLM**), proporcionando retroalimentación casi inmediata y

personalizada. Así, la implementación de **IA** y **PLN** no solo representa una mejora tecnológica, sino también un cambio organizacional que optimiza la gestión de la evaluación docente y la toma de decisiones institucionales. Además, esta innovación se vincula directamente con los **ODS**, especialmente con el **ODS 4** (Educación de Calidad), al elevar la precisión y utilidad de la retroalimentación docente y fomentar prácticas educativas más efectivas, y con el **ODS 9** (Industria, Innovación e Infraestructura), al fortalecer la infraestructura digital y las capacidades institucionales para adoptar tecnologías sostenibles y escalables en educación superior.

Por otro lado, el presente artículo realiza una revisión sistemática de investigaciones publicadas entre 2019 y 2024, con el objetivo de analizar críticamente el uso de ciencia de datos e **IA**, con énfasis en **PLN**, aplicado a la retroalimentación docente personalizada.

Por lo anterior, para guiar este análisis se plantean tres preguntas de investigación: (1) ¿Qué técnicas y algoritmos se reportan para el análisis de datos cuantitativos y cualitativos en la evaluación docente?; (2) ¿Cómo se validan estos modelos y qué nivel de efectividad muestran frente a los métodos tradicionales?; y (3) ¿Cuáles son los beneficios percibidos y los principales desafíos técnicos, éticos e institucionales asociados con su implementación? Las respuestas a estas preguntas se presentan en la **sección 5.1 Síntesis de respuestas** a las preguntas de investigación de este documento. Asimismo, esta revisión sistemática ofrece una síntesis integral que permite mapear las soluciones tecnológicas actuales y evaluar su impacto para modernizar la evaluación docente y promover la mejora continua en el **CUCEA**, constituyendo un aporte relevante a la innovación educativa sostenible.

## **2. CONTEXTUALIZACIÓN**

A continuación, se presenta la contextualización del uso de la **IA** en la evaluación docente universitaria, enmarcando la problemática de los sistemas tradicionales y la transición global hacia soluciones basadas en Ciencia de Datos y **PLN**.

### **2.1. Mundial**

A nivel mundial, el panorama de la educación superior se ha caracterizado históricamente por sistemas de evaluación docente dominados por métodos tradicionales que priorizan promedios

simples y encuestas de preguntas cerradas, estableciendo que estos enfoques resultan insuficientes para capturar la complejidad del desempeño y proporcionar retroalimentación oportuna mientras que, por otro lado, se observa una rápida adopción de la **IA** y el Aprendizaje Automático (**ML**) en el ecosistema educativo para el análisis de sentimientos y la evaluación del desempeño docente (Okoye et al., 2022, 2023).

Ahora bien, las fuentes revelan una tendencia clara hacia la superioridad de los modelos avanzados basados en **PLN** y Aprendizaje Profundo (**DL**) demostrando que técnicas **LLMs** como **BERT** y **DistilBERT**, el análisis de sentimientos y el modelado de tópicos (**LDA**) superan ampliamente a los métodos tradicionales en velocidad, precisión y profundidad analítica para procesar grandes volúmenes de retroalimentación textual (Shuqin & Raga, 2024).

## **2.2. Internacional**

En el ámbito internacional, las instituciones de educación superior funcionan como instituciones productoras de conocimiento debido a que las tendencias globales de evaluación giran en torno a políticas convergentes como el crecimiento de la matrícula, la descentralización administrativa y una mayor competencia por recursos (Martínez-Téllez & Celaya-Lozano, 2023).

En cuanto a la adopción de **IA**, la literatura internacional demuestra un interés creciente, con estudios empíricos en diversas regiones:

- China y Australia han desarrollado modelos híbridos avanzados (como **BERT-BiLSTM-Attention** y **CNN-BiLSTM**) para clasificar emociones complejas en comentarios estudiantiles con alta precisión hasta, 0.9664 **F1-score** (Shuqin & Raga, 2024).
- En Europa y Estados Unidos (**E.U.A.**), se han empleado métodos como **VADER** y análisis léxico para clasificar emociones y habilidades cognitivas (Taxonomía de Bloom), buscando una comprensión más rica del contexto pedagógico (Resuello et al., 2024).
- América Latina, se reporta una actividad emergente liderada por instituciones en Colombia y México, como la Universidad del Valle y la Pontificia Universidad Javeriana, las cuales han comenzado a adoptar **LLMs** y técnicas de **PLN** preentrenados para español —como **Py-sentimiento**—, logrando precisiones del 93% en el análisis de comentarios (Peña-Torres, 2024). No obstante, persiste una brecha investigativa regional marcada por la escasez de datos etiquetados en el idioma local y la fragmentación de los estudios y (Okoye et al., 2022).

### **2.3. País o localidad.**

En México, la evaluación de la enseñanza ha evolucionado mediante herramientas tecnológicas que permiten medir con mayor precisión el desempeño docente y la satisfacción estudiantil (Arriaga-Cárdenas et al., 2022). Se han procesado 471,968 comentarios abiertos en español mediante minería de texto, empleando modelos como **EPDM** para la extracción de sentimientos y valencias emocionales, junto con algoritmos de aprendizaje automático como K vecinos más cercanos (**KNN**) y Máquinas de Vector de Soporte (**SVM**) para predecir recomendaciones docentes basadas en género, emociones y factores pedagógicos integrando métricas estadísticas avanzadas como **MANCOVA**, **ANCOVA**, Kruskal-Wallis, precisión, sensibilidad y curvas (**ROC**), además de modelos de **PLN** que van desde **SVM** hasta **DistilBERT** con precisiones del 93% en español (Okoye et al., 2022, 2023).

Por otro lado, en el **CUCEA**, los sistemas tradicionales de evaluación presentan limitaciones críticas, ya que el procesamiento manual de más de 2'354,631 de registros por ciclo escolar retrasa la entrega de retroalimentación en formato **PDF**, lo que dificulta la toma de decisiones pedagógicas oportunas y evidencia la necesidad de implementar estas soluciones tecnológicas (Santamaría-Velasco, 2024). Por lo anterior, la modernización de los procesos de evaluación requiere no solo herramientas tecnológicas, sino también estrategias para balancear comentarios, evitar sesgos y conformar equipos interdisciplinarios con conocimientos en **IA**, análisis de datos y pedagogía (Vera-Noriega, 2023). Así, el **CUCEA** puede pasar de un análisis superficial a la generación de retroalimentación accionable y personalizada, mejorando la enseñanza y la experiencia estudiantil.

## **3. REVISIÓN DE LA LITERATURA**

La presente sección expone de manera integrada el estado del arte y la base teórica que sustentaron el uso de modelos de Inteligencia Artificial (**IA**) y Procesamiento de Lenguaje Natural (**PLN**) en la evaluación docente universitaria, así como las hipótesis derivadas de la revisión. Para ello se realizó un análisis sistemático de estudios empíricos publicados entre 2019 y 2024, periodo considerado clave para capturar los avances recientes en aprendizaje profundo, modelos transformadores y sistemas de análisis automatizado de texto aplicados a la retroalimentación

educativa, por lo que siguiendo el protocolo **PRISMA**, se revisaron 17 estudios seleccionados de **Scopus**, **Web of Science**, **IEEE Xplore** y **ERIC** donde en la **Tabla 1** se muestra sistematizadamente la relación entre los estudios revisados y su procedencia.

**Tabla 1. Caracterización del corpus de estudio**

Estudio	País	Metodología / Modelo IA	Hallazgo Principal
Pramod et al. (2022)	India	Análisis de sentimientos y Word Cloud en texto y Random Forest como modelo predictivo, destacando modelos de conjunto.	Tras comparar diversos algoritmos (como Naïve Bayes, <b>SVM</b> y Regresión Logística), el modelo RandomForestClassifier fue seleccionado como el más efectivo para predecir la eficacia docente.
Hu et al. (2020)	China	Análisis de sentimientos con <b>SVM</b> (kernels lineal, polinómico, <b>RBF</b> ); preprocesamiento con word2vec, K-means para clustering y <b>TFIDF</b> para ponderar características.	El modelo SVM utilizando un núcleo de base radial ( <b>RBF</b> ) es el método más eficaz y estable para clasificar sentimientos en textos de evaluación docente F1 90.27%
Esparza et al. (2018)	México	Análisis de sentimiento con <b>SVM</b> usando kernels lineal, <b>RBF</b> y polinomial, aprovechando su alto rendimiento en clasificación.	El algoritmo de Máquinas de Vectores de Soporte ( <b>SVM</b> ) con un kernel lineal obtuvo los mejores resultados, logrando una exactitud equilibrada (balanced accuracy) de 0.8149
Peña-Torres et al. (2024)	Colombia	Análisis de sentimiento con <b>NLP</b> usando <b>LLMs</b> preentrenados:Py-sentimiento ( <b>Robertuito/RobERTa</b> ), versión personalizada de <b>VADER</b> y modelo <b>DistilBERT</b> .	El modelo basado en Distilbert alcanzó una métrica del 93% de precisión en la clasificación de la polaridad de los comentarios estudiantiles.
Nikolić et al. (2020)	Serbia	Sistema <b>ABSA</b> con <b>SVM</b> lineal para aspectos y sentimientos, Bag-of-Words con TFIDF, diccionarios y reglas basadas en expresiones regulares.	Se logró un F-measure promedio de 0.90. De manera específica, el sentimiento negativo se detectó con un F-measure de 0.94, mientras que el positivo alcanzó 0.83.
Okoye et al. (2022)	México	Se aplicó el modelo <b>EPDM</b> de minería de texto para extraer sentimientos y valencias emocionales de los estudiantes.	Las métricas de rendimiento del modelo alcanzaron valores óptimos de 1.00 (100%) en precisión, exhaustividad (recall), especificidad, exactitud (accuracy) y puntuación <b>F1</b> , con una tasa de error de 0.00
H. Shuqin & R. C. Raga, (2024)	China	<b>BERT</b> para representaciones contextuales profundas. <b>BiLSTM</b> para captar dependencias bidireccionales. Atención para enfocar la información relevante.	Los modelos superaron en rendimiento a otros enfoques de aprendizaje profundo existentes en el análisis de sentimientos de evaluaciones de cursos. • Exactitud (Accuracy): 0.9608. • Precisión (Precision): 0.9589. • Valor <b>F1 (F1-score)</b> : 0.9664.
Wu et al. (2023)	EEUU	Análisis de sentimiento con <b>VADER</b> .	Al añadir el análisis de sentimiento a la evaluación, la correlación de Pearson con las calificaciones reales de los estudiantes aumentó a 0.16.

Fargues et al. (2023)	Noruega	Se aplicó <b>NLP</b> usando análisis léxico: <b>VADER</b> para sentimientos, <b>NRC</b> para emociones generales y un diccionario personalizado para seis emociones educativas.	Se utilizó una biblioteca de <b>PLN</b> para clasificar el texto en una escala de polaridad: positivo (0.05 a 1), negativo (-0.05 a -1) y neutral (-0.05 a 0.05)
Peng et al. (2022)	China	Modelo de análisis de sentimientos que combina <b>CNN</b> , <b>BLSTM</b> y atención, usando Softmax para clasificar emociones complejas y optimizado con descenso de gradiente mini-batch.	El modelo propuesto obtuvo un <b>F1</b> de 0.845 y un mínimo de 0.748 en los mecanismos de atención.
Shaik et al. (2022)	Australia	<b>IA</b> y <b>PNL</b> : preprocesamiento, TextBlob para sentimiento, Word2Vec para vectores y un modelo <b>Bi-LSTM</b> para clasificar comentarios en 19 subdominios de Biggs.	En estudios relacionados citados, se reportó un <b>F1-score</b> de 86.13% para la identificación de categorías de aspectos y un 82.10% para su clasificación de sentimiento.
Pu et al. (2021)	China	Modelo que combina Word2Vec y GloVe para representar palabras, usa <b>CNN</b> y <b>BiLSTM</b> para extraer características, y optimiza un ensemble con el algoritmo <b>MOGWO</b> .	Se obtuvo un <b>F1</b> superior al 91% (específicamente 0.9184), superando a otros siete modelos de comparación como <b>SVM</b> , <b>RNN</b> y <b>CNN</b> tradicionales.
Resuello et al. (2024)	Filipinas	Se aplicaron técnicas de minería de texto, <b>NLP</b> , <b>VADER</b> y <b>LDA</b> .	Se alcanzó un grado de acuerdo del 70.14% entre <b>VADER</b> y <b>TextBlob</b> en los comentarios negativos
Okoye et al. (2023)	México	Se aplicó el modelo <b>EPDM</b> de minería de texto para extraer sentimientos y valencias emocionales de los estudiantes.	para cuantificar datos textuales, detectando que la confianza (~25%) fue la emoción predominante en las evaluaciones a instructores, mientras que la preocupación (~20%) destacó en las recomendaciones entre estudiantes.
Das et al. (2022)	India	Regresión Logística, <b>XGBoost</b> y Árboles de Decisión.	El valor más alto alcanzado fue un <b>F1-score</b> de 0.728 mediante la combinación de Regresión Logística y <b>TF-IDF</b> .
Yuan & Hernandez, (2023)	Filipinas	<b>Bi-LSTM</b> .	Las métricas clave obtenidas por este modelo son: <ul style="list-style-type: none"> <li>• Exactitud de validación: Alcanzó un 90.77%,.</li> <li>• Exactitud de prueba: Logró un 84.33%,.</li> </ul>
Ponmalar et al. (2024)	India	<b>PLN</b> y <b>BERT</b> .	Se evaluaron las respuestas descriptivas en educación en línea con una precisión promedio del 90.6% (en un rango del 87% al 95%)

Fuente: Elaboración propia basada en la caracterización del corpus de estudio.

Ahora bien, el análisis bibliométrico mostró un predominio de investigaciones orientadas a métricas técnicas (**precisión** y **F1-score**), con un desarrollo acelerado en Asia y Norteamérica,

mientras que los estudios en contextos hispanohablantes permanecieron limitados, fragmentados y con escasez de corpus etiquetados en español. Por otro lado, la exclusión de artículos en idiomas distintos al inglés y español constituyó un sesgo recurrente, y resaltó la limitada diversidad lingüística de la literatura disponible.

Históricamente, la evaluación docente se fundamentó en instrumentos tradicionales basados en encuestas cerradas y promedios simples, enfoques que resultaron insuficientes para capturar la complejidad del desempeño y la diversidad de percepciones estudiantiles (Shaik et al., 2022). En contraste, el estado del arte mostró que las técnicas modernas de **PLN** superaron estos métodos en precisión, velocidad y profundidad analítica, ya que modelos clásicos como Máquinas de Vector de Soporte (**SVM**) se aplicaron en estudios Mexicanos para clasificar comentarios en español, aunque su desempeño varió según el kernel utilizado (Okoye et al., 2022).

Sin embargo, los métodos basados en aprendizaje profundo demostraron superioridad: por ejemplo, el modelo híbrido representaciones contextuales bidireccionales con redes de memoria a largo y corto plazo y mecanismos de atención (**BERT-BiLSTM-Attention**) en China alcanzó la métrica **F1-score** de 0.9664, y el uso de las Redes de Memoria a Largo y Corto Plazo Bidireccionales (**Bi-LSTM**) en Filipinas superó el 90% de precisión en tendencias emocionales, donde particularmente relevantes fueron los Modelos de Lenguaje de Gran Escala **LLMs** adaptados al español.

Asimismo, en un estudio comparativo, la versión destilada y multilingüe de **BERT multilingual-DistilBERT** logró un 93% de precisión, superando ampliamente a enfoques léxicos como el Diccionario y Razonador de Sentimientos Consciente de la Valencia (**VADER**) de 0.79, lo que evidenció su capacidad predictiva y comprensión contextual avanzada (Shuqin & Raga, 2024).

Por otro lado, la literatura también destacó herramientas de mayor profundidad analítica, por ejemplo, el modelado de tópicos mediante la Asignación Latente de Dirichlet (**LDA**) se utilizó para vincular temas subyacentes en grandes volúmenes de texto cualitativo, mientras que el Análisis de Sentimiento Basado en Aspectos (**ABSA**) permitió detectar simultáneamente el aspecto educativo mencionado y la polaridad asociada, ofreciendo un entendimiento más fino de la percepción estudiantil (Nikolić et al., 2020).

Otros trabajos clasificaron comentarios según la **Taxonomía de Bloom**, ampliando el valor pedagógico del análisis cualitativo (Fargues et al., 2023). Asimismo, modelos predictivos integrados en sistemas de soporte a la decisión, como el Modelo de Descubrimiento de Parámetros Emocionales más Aprendizaje Automático (**EPDM+ML**), mostraron capacidad para anticipar niveles de satisfacción estudiantil y apoyar procesos de mejora institucional (Okoye et al., 2022). A pesar de sus ventajas, el estado del arte señaló desafíos relevantes, una baja interpretabilidad de los modelos de aprendizaje profundo, la presencia de sesgos algorítmicos (por ejemplo, de género), la resistencia institucional a la adopción tecnológica y la ausencia de validaciones longitudinales que evaluaran el impacto sostenido de estas herramientas (**LDA, ABSA y EPDM+ML**). Además, se subrayó también la necesidad de complementar los modelos con validación humana para fortalecer su legitimidad y utilidad educativa (Peña-Torres, 2024).

Por lo tanto, a partir del análisis sistemático de la literatura, se identifican patrones de rendimiento que permiten formular los supuestos de esta investigación. En la **Tabla 1** se sistematiza los **17 estudios que conforman el corpus de esta revisión**, detallando el modelo utilizado y su precisión, lo que sirve de base empírica para las **hipótesis** planteadas a continuación:

**H1:** *“Basado en los resultados de Shuqin & Raga (2024) y la comparativa de modelos multilingües analizada en la **Tabla 1**, se postula que los modelos de aprendizaje profundo basados en transformadores (como **DistilBERT**) adaptados al español, ofrecen una precisión analítica superior al 90% en la clasificación de sentimientos, superando significativamente la capacidad de los métodos léxicos tradicionales como **VADER** (0.79).”*

**H2:** *“A partir de la capacidad demostrada por las técnicas de **ABSA** y **LDA** para desglosar dimensiones cualitativas (Nikolić et al., 2020), se asume que su integración permite identificar no solo la polaridad, sino los aspectos específicos del proceso educativo, generando una retroalimentación más detallada y accionable que los promedios numéricos simples.”*

**H3:** *“Considerando que los modelos de **IA / PLN** reducen los tiempos de procesamiento en grandes volúmenes de texto (Okoye et al., 2022), se hipotetiza que su implementación en el contexto del **CUCEA** permitirá procesar de forma automatizada los más de 2'354,631 de registros por ciclo, transformando datos masivos en decisiones oportunas para la mejora docente.”*

Donde para el análisis de las hipótesis planteadas, se establecieron criterios analíticos de comparación basados exclusivamente en la revisión sistemática reportada y el análisis comparativos de la literatura. En el caso de la **H1**, la comparación se sustenta en la precisión analítica reportada y en el desempeño relativo de los modelos de aprendizaje profundo basados en transformadores frente a métodos léxicos tradicionales. Por el contrario, en la **H2** se analiza a partir del nivel de granularidad del análisis, considerando la capacidad de distintas técnicas para identificar dimensiones cualitativas y aspectos específicos más allá de la polaridad global. Finalmente, la **H3** se fundamenta en la capacidad de procesamiento y escalabilidad de los modelos de **IA** y **PLN**, evaluando su potencial para automatizar el análisis de grandes volúmenes de texto y apoyar la toma de decisiones oportunas en contextos institucionales, sin introducir nuevas variables ni mediciones propias.

#### **4. METODOLOGÍA**

Esta revisión sistemática se diseñó siguiendo un enfoque riguroso y transparente, con el objetivo de identificar, seleccionar, analizar e interpretar estudios relevantes sobre el uso de técnicas de **IA**, particularmente aquellas basadas en **NLP**, para la retroalimentación docente en el contexto de la educación superior, por lo tanto, a continuación, se describen los métodos empleados para garantizar la validez y reproducibilidad del proceso.

##### **4.1 Estrategia de búsqueda**

Se definió una estrategia de búsqueda estructurada mediante operadores Booleanos y de proximidad, con el fin de capturar de manera precisa los estudios más relevantes, por lo que las cadenas de búsqueda fueron diseñadas para cubrir cuatro dimensiones clave: Evaluación docente, técnicas de **PLN**, métodos de **IA** y contexto universitario. Un ejemplo de cadena utilizada es:

("teacher evaluation" OR "faculty evaluation" OR "teaching feedback") AND ("natural language processing" OR "text mining" OR "sentiment analysis") AND ("machine learning" OR "artificial intelligence") AND ("higher education" OR "university").

Estas búsquedas se llevaron a cabo en bases de datos académicas como: **Scopus**, **Web of Science**, **IEEE Xplore** y **ERIC**, considerando un rango temporal de 2019 a 2024, con el objetivo

de capturar tanto los desarrollos más recientes como estudios relevantes que sirvan como antecedentes clave en el área.

## **4.2. Criterios**

Por otro lado, se establecieron criterios claros de inclusión y exclusión para asegurar la pertinencia y calidad de los estudios seleccionados: Criterios de inclusión se incluirán estudios empíricos (cuantitativos, cualitativos o mixtos) que describan el desarrollo, implementación o evaluación de modelos o sistemas de **IA** aplicados específicamente a la retroalimentación docente en contextos universitarios y criterios de exclusión: Se excluirán artículos que no estén escritos en inglés o español, estudios enfocados en educación primaria o secundaria, artículos de opinión, revisiones teóricas o sin datos empíricos, así como modelos aplicados a otros ámbitos que no involucren directamente la evaluación docente universitaria.

## **4.3. Fases**

El proceso de selección se realizó en dos fases sucesivas, primero, se examinaron los títulos y resúmenes de los estudios recuperados para filtrar aquellos que cumplen con los criterios de elegibilidad. En la segunda fase, se analizó el texto completo de los estudios preseleccionados realizándose de forma independiente por dos revisores, donde cualquier discrepancia fue resuelta mediante consenso o la intervención de un tercer revisor. Además se documentó el proceso completo utilizando el diagrama de flujo Preferred Reporting Items for Systematic Reviews and Meta-Analyses (**PRISMA** por sus siglas en inglés) ver **Figura 1**.

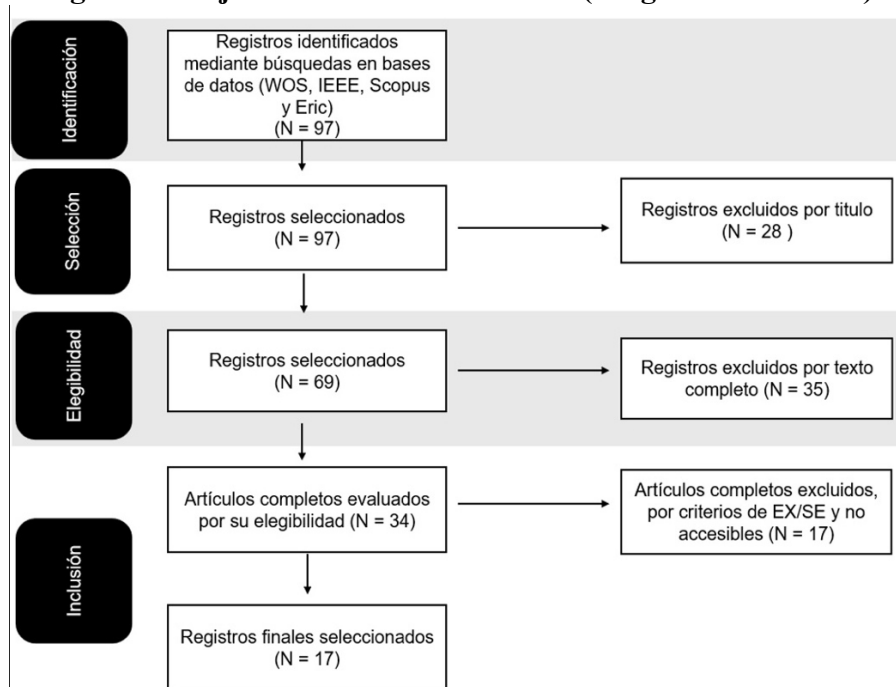
Para la extracción de datos se diseñó una hoja de recolección estandarizada que incluye las siguientes variables: Autor(es), año de publicación, país de origen del estudio, objetivo principal, técnicas de **IA** utilizadas, tipo de datos analizados (cuantitativos y/o cualitativos), métricas de validación empleadas (tanto técnicas como de percepción), resultados principales, limitaciones reportadas y desafíos identificados. Asimismo, el análisis de los datos fue de carácter narrativo y temático, guiado por las preguntas de investigación previamente planteadas, siendo los hallazgos agrupados en torno a las siguientes categorías: Técnicas y algoritmos utilizados, métodos de validación y efectividad, beneficios percibidos, y barreras técnicas, éticas e institucionales para la

implementación de estas tecnologías en entornos universitarios, presentándose en la **Tabla 2** de la siguiente sección.

## 5. RESULTADOS

El proceso de selección se desarrolló bajo los lineamientos del diagrama **PRISMA**, partiendo de 97 registros identificados en **Web of Science, IEEE Xplore, Scopus y ERIC**. Así mismo para la fase de cribado o selección fue realizada por dos revisores de manera independiente, donde las discrepancias surgidas durante este proceso se resolvieron mediante la discusión técnica y el consenso mutuo, garantizando así la rigurosidad en la aplicación de los criterios. No se aplicó el estadístico Kappa de Cohen debido a que el diseño metodológico priorizó la validación cualitativa por consenso sobre la medición de la concordancia aleatoria, asegurando que cada inclusión fuera validada por la totalidad del equipo investigador. Tras excluir 28 registros por título y 35 por criterios de inclusión en texto completo, se evaluaron 34 artículos finales. De estos, 17 fueron descartados por incumplimiento de criterios específicos o inaccesibilidad, resultando en 17 estudios incluidos. La **Figura 1** detalla este flujo de decisiones mediante el diagrama de flujo **PRISMA**.

**Figura 1. Flujo de selección de estudios (Diagrama PRISMA)**



Fuente: Elaboración propia basada en el proceso de selección conforme a los lineamientos del diagrama PRISMA.

Ahora bien, con el fin de organizar y presentar de manera clara la información recopilada en esta revisión, se ha elaborado una tabla, ver **Tabla 2**, la cual resume distintos aspectos clave de los estudios analizados permitiendo visualizar de forma estructurada los datos más relevantes como: País de origen, objetivo del estudio, técnicas de IA utilizadas, tipo de datos de entrada, métricas de validación, resultados principales, limitaciones reportadas y desafíos identificados relevantes para esta revisión.

**Tabla 2. Principales características de los estudios revisados**

Estudio	País	Objetivo del estudio	Técnicas de IA utilizadas	Tipo de datos de entrada
Pramod et al. (2022)	India	Predecir la efectividad docente mediante machine learning.	Análisis de sentimientos y Word Cloud en texto y Random Forest como modelo predictivo, destacando modelos de conjunto.	Análisis de datos cuantitativos y cualitativos de cuestionarios estudiantiles.
Hu et al. (2020)	China	Proponer un método de análisis de sentimientos aplicando machine learning.	Análisis de sentimientos con SVM (kernels lineal, polinómico, RBF); preprocesamiento con word2vec, K-means para clustering y TFIDF para ponderar características.	Datos cualitativos: 581 evaluaciones cortas en chino, 566 usadas como corpus (466 entrenamiento, 100 prueba), con textos de 25 a 150 palabras.
Esparza et al. (2018)	México	Clasificar comentarios reales en español.	Análisis de sentimiento con SVM usando kernels lineal, RBF y polinomial, aprovechando su alto rendimiento en clasificación.	Datos cualitativos: 1040 comentarios en español de estudiantes de ingeniería, evaluando 21 profesores; 99 características identificadas tras limpieza.
Peña-Torres et al. (2024)	Colombia	Mejorar la práctica docente.	Análisis de sentimiento con NLP usando LLMs preentrenados: Py-sentimiento (Robertuito/RobERTa), versión personalizada de VADER y modelo DistilBERT.	Enfoque mixto: cuantitativo (medidas de rendimiento LLM) y cualitativo (caso de estudio, prototipo). Datos: 365 comentarios en español.
Nikolić et al. (2020)	Serbia	Analizar sentimientos en texto libre de encuestas estudiantiles en serbio.	Sistema ABSA con SVM lineal para aspectos y sentimientos, Bag-of-Words con TFIDF,	Datos cualitativos en serbio: Presencial (2,472 reseñas, 2,798 oraciones) y C2 online

Estudio	País	Objetivo del estudio	Técnicas de IA utilizadas	Tipo de datos de entrada
			diccionarios y reglas basadas en expresiones regulares.	(3,863 reseñas y 6,896 oraciones).
Okoye et al. (2022)	México	Evaluar la evaluación de del conocimiento adquirido por los estudiantes.	Se aplicó el modelo <b>EPDM</b> de minería de texto para extraer sentimientos y valencias emocionales de los estudiantes.	Cualitativamente se cuantificaron 471,968 comentarios abiertos en español.
Shuqin & Raga (2024)	China	Proponer el modelo <b>BERTBiLSTM-ATTENTION (BBA)</b> para extraer sentimientos de evaluaciones de cursos.	<b>BERT</b> para representaciones contextuales profundas. <b>BiLSTM</b> para captar dependencias bidireccionales. Atención para enfocar la información relevante.	Se analizaron 9944 comentarios reales de estudiantes de cursos de informática en la plataforma <b>MOOC</b> de una universidad China.
Wu et al (2023)	EEUU	Mejorar resultados de aprendizaje y rediseñar el curso <b>AME4163</b> .	Análisis de sentimiento con <b>VADER</b> .	Textos no estructurados de estudiantes en forma de take-aways ( <b>TA</b> ) y oraciones de aprendizaje ( <b>LS</b> ) estructuradas como tripletas (Experiencia, Aprendizaje y Valor).
Fargues et al. (2023)	Noruega	Analizar la retroalimentación abierta de estudiantes.	Se aplicó <b>NLP</b> usando análisis léxico: <b>VADER</b> para sentimientos, <b>NRC</b> para emociones generales y un diccionario personalizado para seis emociones educativas.	Los datos de entrada son cualitativos, provenientes de retroalimentación abierta de estudiantes.
Peng et al. (2022)	China	Desarrollar un modelo de análisis de sentimientos aplicado a las evaluaciones docentes.	Modelo de análisis de sentimientos que combina <b>CNN</b> , <b>BLSTM</b> y atención, usando Softmax para clasificar emociones complejas y optimizado con descenso de gradiente mini-batch.	Datos cuantitativos (precisión, pérdida, recall, <b>F1-score</b> ) y cualitativos (comentarios de estudiantes con carga emocional).
Shaik et al. (2022)	Australia	Desarrollar una aplicación que permita a una institución educativa analizar comentarios textuales para evaluar y mejorar	<b>IA y PNL</b> : preprocesamiento, <b>TextBlob</b> para sentimiento, <b>Word2Vec</b> para vectores y un modelo <b>Bi-LSTM</b>	El estudio analizó 5,585 comentarios cualitativos y cuantitativos de estudiantes (texto y calificaciones de 1 a 5)

Estudio	País	Objetivo del estudio	Técnicas de IA utilizadas	Tipo de datos de entrada
		sus sistemas y servicios.	para clasificar comentarios en 19 subdominios de Biggs.	del semestre 2 de 2021.
Pu et al. (2021)	China	Proponer un modelo de aprendizaje profundo en conjunto para analizar sentimientos en evaluaciones de cursos online.	Modelo que combina <b>Word2Vec</b> y <b>GloVe</b> para representar palabras, usa <b>CNN</b> y <b>BiLSTM</b> para extraer características, y optimiza un ensemble con el algoritmo <b>MOGWO</b> .	El conjunto de datos preprocesado incluye 20,240 entradas con etiquetas emocionales.
Resuello et al. (2024)	Filipinas	Utilizar algoritmos de minería de texto para analizar comentarios estudiantiles sobre desempeño docente.	Se aplicaron técnicas de minería de texto, <b>NLP</b> , <b>VADER</b> y <b>LDA</b> .	Se analizaron principalmente comentarios cualitativos de estudiantes, con un total de 2,312 comentarios recopilados entre 2020 y 2023.
Okoye et al. (2023)	México	El estudio analizó la evaluación de la enseñanza por parte de los estudiantes ( <b>SET</b> ) para identificar los factores clave que afectan el desempeño y la enseñanza en la Tecnología.	Se aplicó el modelo <b>EPDM</b> de minería de texto para extraer sentimientos y valencias emocionales de los estudiantes.	Enfoque mixto (cualitativo y cuantitativo): Cualitativamente se cuantificaron 471,968 comentarios abiertos en español.
Das et al. (2022)	India	Mejorar la enseñanza usando <b>ML</b> .	Regresión Logística, <b>XGBoost</b> y Árboles de Decisión.	El estudio se enfoca en la Retroalimentación Cualitativa ( <b>QLF</b> ).
Yuan & Hernandez (2023)	Filipinas	Analizar la tendencia emocional de los estudiantes.	<b>Bi-LSTM</b> .	Textos de evaluación docente de estudiantes obtenidos mediante evaluación en línea.
Ponmalar et al. (2024)	India	Proponer el AI-Enhanced Context-Aware Semantic Grading Network ( <b>AI-CASGN</b> ).	<b>PLN</b> y <b>BERT</b> .	Evaluación de respuestas descriptivas (datos cualitativos, texto), que requieren una visión más profunda de la comprensión, la capacidad de razonamiento y el pensamiento analítico de los estudiantes.

Fuente: Elaboración propia basada en las principales características de los estudios revisados

Por otro lado, respecto a la primera interrogante planteada en la introducción, la literatura científica revisada evidencia una transición tecnológica clara, donde el PLN ha evolucionado desde el análisis de sentimientos basado en diccionarios léxicos y modelos de aprendizaje automático tradicional, como el modelo SVM, hacia arquitecturas de aprendizaje profundo altamente sofisticadas. En este sentido, el uso predominante de modelos como las redes de memoria a largo y corto plazo bidireccionales y los transformadores tipo BERT ha permitido no solo clasificar polaridades, sino también identificar temas subyacentes mediante el modelado de tópicos con la técnica de Asignación Latente de Dirichlet (LDA) y categorizar el nivel de pensamiento crítico a través de la Taxonomía de Bloom.

En lo que respecta a la segunda pregunta de investigación, los datos indican que los modelos avanzados como DistilBERT logran una efectividad superior a los métodos tradicionales, alcanzando precisiones cercanas al 93% en contraste con el 79% que suelen reportar herramientas léxicas como VADER. Esta superioridad técnica se valida mediante métricas estandarizadas como el F1-score y el área bajo la curva, permitiendo además una eficiencia operativa notable al reducir los tiempos de respuesta a escasos 3.2 segundos.

En cuanto a la tercera pregunta de investigación, sobre los beneficios y desafíos, la investigación revela que la automatización analítica permite generar retroalimentación personalizada y escalable para grandes volúmenes de datos, como los más de 2'354,631 de registros generados por ciclo en el CUCEA. No obstante, este potencial se enfrenta a retos éticos y técnicos significativos, tales como la baja interpretabilidad de los modelos de "caja negra", la presencia de sesgos algorítmicos de género y la resistencia cultural al cambio tecnológico dentro del sector educativo. Atendiendo a estas implicaciones.

## **6. DISCUSIÓN**

Desde una perspectiva institucional, los resultados evidencian consistencia entre los planteamientos teóricos formulados y los patrones identificados en la literatura especializada. En este sentido, las hipótesis (H1, H2 y H3) encuentran respaldo conceptual en la evidencia sintetizada, al observarse convergencias significativas en los estudios analizados. Dicho respaldo se fundamenta en análisis comparativo y revisión sistemática, y no en validación experimental directa ni en datos primarios generados por la investigación. Además es importante señalar, que la

principal contribución del artículo es de carácter metodológico-organizacional, al proponer un marco replicable de evaluación docente basado en **IA/PLN** como innovación de proceso, más que el desarrollo de un algoritmo específico. Casos como los centros universitarios de la Red Universitaria de la Universidad de Guadalajara —entre ellos el Centro Universitario de Ciencias Económico-Administrativas el cual se presenta como un contexto institucional ilustrativo, y no como un estudio empírico aplicado. — ilustran desafíos comunes a numerosas instituciones latinoamericanas, donde el procesamiento de millones de registros por periodo limita la oportunidad y utilidad de la retroalimentación.

En este contexto, la literatura revisada evidencia que la integración de **IA**, ciencia de datos y **PLN** constituye una innovación de proceso organizacional, en los términos establecidos por el Manual de Oslo, al automatizar el análisis de información cualitativa a gran escala y mejorar la eficiencia, calidad y trazabilidad de los procesos evaluativos. Modelos como **AI-CASGN**, capaces de generar retroalimentación en segundos, y arquitecturas basadas en **DistilBERT**, que alcanzan altos niveles de precisión en español, demuestran la superioridad de estos enfoques frente a métodos tradicionales y su viabilidad para entornos universitarios latinoamericanos.

Asimismo, el uso de modelos híbridos como **BERT-BiLSTM-ATTENTION** y técnicas de modelado de tópicos permite capturar tanto la semántica como las sutilezas emocionales presentes en los comentarios estudiantiles, favoreciendo una comprensión más profunda de la experiencia educativa y fortaleciendo los mecanismos de mejora continua del desempeño docente.

Desde la perspectiva del desarrollo sostenible, los hallazgos de este estudio contribuyen directamente al fortalecimiento del **ODS 4**, al proponer mecanismos innovadores que mejoran la calidad de la educación superior mediante procesos de evaluación docente más justos, oportunos y formativos. De igual manera, se alinean con el **ODS 9** al promover el uso estratégico de infraestructura digital y tecnologías avanzadas para optimizar procesos organizacionales en instituciones educativas.

La adopción de modelos de **IA** y **PLN** no solo incrementa la eficiencia operativa de los sistemas de evaluación, sino que también favorece una gestión del conocimiento más sostenible, al transformar grandes volúmenes de datos en información útil para la toma de decisiones, el diseño de políticas académicas y la mejora institucional en universidades latinoamericanas.

A diferencia de revisiones previas centradas exclusivamente en métricas técnicas o en aplicaciones aisladas de IA en educación, este estudio integra de manera sistemática la evaluación docente universitaria, la Inteligencia Artificial y el Procesamiento de Lenguaje Natural desde un enfoque de innovación de proceso organizacional, alineado con el Manual de Oslo y los Objetivos de Desarrollo Sostenible. Esta aproximación permite articular rigor metodológico, análisis técnico e implicaciones institucionales, ofreciendo un marco analítico replicable para redes universitarias y universidades públicas de América Latina.

### **6.1 Implicaciones Teóricas (Scientia)**

La presente investigación realiza una contribución significativa al estado del conocimiento al proponer un enfoque interdisciplinario que trasciende la visión puramente técnica de la Inteligencia Artificial (IA), esto debido a que se diferencia de revisiones previas que abordan el análisis de sentimientos o la minería de datos de forma aislada. Asimismo, este estudio integra el uso de la IA y el Procesamiento de Lenguaje Natural (PLN) dentro de un marco que vincula la valuación docente con la innovación de proceso, bajo las directrices del Manual de Oslo, y los Objetivos de Desarrollo Sostenible (ODS 4 y 9).

Desde una perspectiva teórica, los hallazgos subrayan dos vacíos críticos en la literatura de habla hispana:

- **Contextualización Lingüística y Cultural:** Se identifica que, aunque modelos multilingües como **DistilBERT** alcanzan precisiones del 93%, persiste una escasez de corpus etiquetados en español que capturen los matices educativos regionales. Esta investigación aporta una base teórica para la transición de modelos léxicos hacia arquitecturas de aprendizaje profundo adaptadas.
- **Validez de Constructo y Medición Longitudinal:** Se plantea la necesidad de marcos conceptuales que no solo midan la polaridad (positivo/negativo), sino que integren dimensiones pedagógicas complejas como la Taxonomía de Bloom, ya que la literatura actual carece de evidencia longitudinal, por lo que este estudio fundamenta teóricamente cómo la retroalimentación automatizada debe evaluarse en función de la mejora sostenida de las estrategias docentes y la satisfacción estudiantil a largo plazo.

### **6.2 Implicaciones Prácticas**

En términos de aplicación inmediata, este artículo trasciende la teoría al proponer un diseño de plataforma concreto para el contexto del Centro Universitario de Ciencias Económico Administrativas (CUCEA), ya que esta propuesta aborda directamente las limitaciones críticas de baja precisión, análisis superficial y retroalimentación tardía de los sistemas actuales a través de las siguientes dimensiones:

- **Eficiencia Operativa y Escalabilidad:** La integración de Modelos de Lenguaje de Gran Escala (LLMs) permite el procesamiento ágil de grandes volúmenes de datos, específicamente los más de 2'354,631 registros generados por ciclo en el CUCEA, garantizando una capacidad de respuesta que los métodos manuales o léxicos tradicionales no pueden alcanzar.
- **Profundidad Analítica (Accionabilidad):** Mediante la implementación de técnicas de Análisis de Sentimiento Basado en Aspectos (ABSA) y Asignación Latente de Dirichlet (LDA), la plataforma transforma comentarios abiertos en datos estructurados por dimensiones (metodología, evaluación, infraestructura). Esto permite a los gestores académicos y docentes recibir retroalimentación predictiva y en tiempo real a través de Sistemas de Soporte a la Decisión (DSS).
- **Ética y Gobernanza de Datos:** La implementación práctica exige un diseño centrado en la transparencia algorítmica, por lo que se proponen mecanismos específicos para mitigar sesgos de género y asegurar la privacidad de los datos, elementos clave para vencer la resistencia institucional y fomentar la adopción activa de la tecnología por parte del profesorado.

## 7. CONCLUSIÓN

El presente artículo de **revisión sistemática (2019–2024)** confirma la **insuficiencia de los modelos tradicionales de evaluación docente**, caracterizados por el uso de promedios simples y análisis manuales que generan una retroalimentación superficial, tardía y altamente subjetiva. Esta situación limita la toma de decisiones informadas para la mejora continua de la docencia y resulta especialmente crítica en contextos universitarios que enfrentan procesos de expansión, diversificación de modalidades educativas y mayores exigencias de calidad académica.

La evidencia analizada demuestra que la **incorporación de técnicas de Inteligencia Artificial**, particularmente aquellas orientadas al análisis automatizado de datos cualitativos y cuantitativos,

ofrece una **oportunidad estratégica para transformar los sistemas de evaluación docente**. Estos enfoques permiten identificar patrones complejos, generar retroalimentación oportuna y personalizada, y fortalecer una cultura de evaluación formativa centrada en el desarrollo profesional del docente, más allá del carácter meramente administrativo o sancionador.

Desde una el punto de vista prospectivo, los hallazgos de esta revisión abren la posibilidad de **desarrollar estudios longitudinales** que permitan analizar la evolución del desempeño docente a lo largo del tiempo, evaluar el impacto sostenido de la retroalimentación inteligente y medir su influencia en variables clave como la calidad del aprendizaje, la innovación pedagógica y la satisfacción estudiantil. Este tipo de estudios resultan fundamentales para transitar de diagnósticos estáticos hacia modelos dinámicos de mejora continua basados en evidencia.

Asimismo, los resultados sugieren un **alto potencial de replicabilidad del modelo en universidades latinoamericanas**, dado que muchas instituciones comparten problemáticas estructurales similares: limitaciones tecnológicas, sobrecarga administrativa y sistemas de evaluación poco contextualizados. La adaptación de soluciones basadas en **IA**, considerando marcos normativos, realidades institucionales y enfoques de justicia educativa propios de la región, puede contribuir significativamente a la modernización de la gestión académica y al fortalecimiento de la calidad educativa en América Latina.

En síntesis, esta revisión no solo evidencia la necesidad de superar los modelos tradicionales de evaluación docente, sino que también **proyecta una línea clara de investigación y acción institucional**, orientada a la implementación de sistemas inteligentes, éticos y contextualizados, capaces de responder a los desafíos presentes y futuros de la educación superior.

Finalmente, como trabajo futuro, se propone la validación aplicada del enfoque presentado a través de la implementación un estudio empírico posterior en un contexto educativo real.

## 8. REFERENCIAS

- Arriaga-Cárdenas, O. G., Lara-Magaña, P. del C., & Pasciuta-Marco, P. D. (2022). Los centros públicos de investigación, como eje central de la innovación y la educación en México. *Revista científica Scientia et PRAXIS*, 2(04), 66–81. <https://doi.org/10.55965/setp.2.coed.a3>
- Das, S., Roy, S., Bose, R., Acharjya, P. P., & Mondal, H. (2022). Analysis of student sentiment dynamics to evaluate teachers' performance in online courses using machine learning. In *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)* (pp. 668–673). IEEE. <https://doi.org/10.1109/ICAAIC53929.2022.9792958>

- Esparza, G. G., de-Luna, A., Zezzatti, A. O., Hernandez, A., Ponce, J., Álvarez, M., Cossio, E., & de Jesus Nava, J. (2018). A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. In *14th International Conference Distributed Computing and Artificial Intelligence* (Vol. 14, pp. 157–164). Springer Nature. [https://doi.org/10.1007/978-3-319-62410-5\\_19](https://doi.org/10.1007/978-3-319-62410-5_19)
- Fargues, M., Kadry, S., Lawal, I. A., Yassine, S., & Rauf, H. T. (2023). Automated analysis of open-ended students' feedback using sentiment, emotion, and cognition classifications. *Applied Sciences*, *13*(4), 2061. <https://doi.org/10.3390/app13042061>
- Hu, X., Yang, Y., Wu, X., & Li, Y. (2020). Text analysis of teaching evaluation based on machine learning. In *Proceedings of the 10th International Workshop on Computer Science and Engineering (WCSE 2020)*. WCSE. <https://doi.org/10.18178/wcse.2020.06.004>
- Martínez-Téllez, M. Á., & Celaya-Lozano, A. (2023). Evaluación institucional de contenidos y proyectos de investigación en un CPI: Caso de estudio (2013-2022). *Scientia et PRAXIS*, *3*(05), 51–77. <https://doi.org/10.55965/setp.3.coed1.a3>
- Nikolić, N., Grljević, O., & Kovačević, A. (2020). Aspect-based sentiment analysis of reviews in the domain of higher education. *The Electronic Library*, *38*(1), 44–64. <https://doi.org/10.1108/EL-06-2019-0140>
- Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Achem, J. A. G., Escamilla, J., & Hosseini, S. (2022). Towards teaching analytics: A contextual model for analysis of students' evaluation of teaching through text mining and machine learning classification. *Education and Information Technologies*, *27*(3), 3891–3933. <https://doi.org/10.1007/s10639-021-10751-5>
- Okoye, K., Daruich, S. D. N., de La O, J. F. E., Castano, R., Escamilla, J., & Hosseini, S. (2023). A text mining and statistical approach for assessment of pedagogical impact of students' evaluation of teaching and learning outcome in education. *IEEE Access*, *11*, 9577–9596. <https://doi.org/10.1109/ACCESS.2023.3239779>
- Peng, H., Zhang, Z., & Liu, H. (2022). A sentiment analysis method for teaching evaluation texts using attention mechanism combined with CNN-BLSTM model. *Scientific Programming*, *2022*, 1–9. <https://doi.org/10.1155/2022/8496151>
- Peña-Torres, J. A. (2024). Towards an improved of teaching practice using sentiment analysis in student evaluation. *Ingeniería y Competitividad*, *26*(2). <https://doi.org/10.25100/iyc.v26i2.13759>
- Ponmalar, S. P., V. S., Srithar, V., Morankar, G., Jebaseelan, S. D. S., & Harikrishnan, R. (2024). *AI-driven CASGN framework for evaluating descriptive answers in online education*. In *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1–8). IEEE. <https://doi.org/10.1109/ICSES63760.2024.10910723>
- Pramod, D., Bharathi, S. V., & Raman, R. (2022). Faculty effectiveness prediction using machine learning and text analytics. In *2022 IEEE Technology and Engineering Management Conference (TEMSCON EUROPE)* (pp. 40–47). IEEE. <https://doi.org/10.1109/TEMSCONEUROPE54743.2022.9801997>
- Pu, X., Yan, G., Yu, C., Mi, X., & Yu, C. (2021). Sentiment analysis of online course evaluation based on a new ensemble deep learning mode: Evidence from Chinese. *Applied Sciences*, *11*(23), 11313. <https://doi.org/10.3390/app112311313>

- Resuello, J. L., Feliscuzo, L. S., & Pantaleon, C. B. (2024). Text mining student comments for teaching performance evaluation using VADER and Latent Dirichlet Allocation algorithm. In *2024 International Conference Automatics and Informatics (ICAI)* (pp. 260–265). IEEE. <https://doi.org/10.1109/ICAI63388.2024.10851580>
- Santamaria-Velasco, C. A., Montiel-Méndez, O. J. M.-M., & Montañez-Moya, G. S. (2024). Liderazgo transformacional y emprendimiento en estudiantes: Una vía hacia el desarrollo educativo sostenible. *Scientia et PRAXIS*, 4(08), 90–119. <https://doi.org/10.55965/setp.4.08.uady.a4>
- Shaik, T., Tao, X., Dann, C., Quadrelli, C., Li, Y., & O'Neill, S. (2022). Educational decision support system adopting sentiment analysis on student feedback. In *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (pp. 377–383). IEEE. <https://doi.org/10.1109/WI-IAT55865.2022.00062>
- Shuqin, H., & Raga, R. C. (2024). A deep learning model for student sentiment analysis on course reviews. *IEEE Access*, 12(2169–3536), 136747–136758. <https://doi.org/10.1109/ACCESS.2024.3463793>
- Vera-Noriega, J. Á., Borbón-Morales, C. G., Mejía-Trejo, J., & Durazo-Salas, F. F. (2023). Relación y comparación entre las variables de clima organizacional, satisfacción y calidad de vida en una empresa de ventas por teléfono en el noroeste de México. *Scientia et PRAXIS*, 3(05), 83–109. <https://doi.org/10.55965/setp.3.05.a4>
- Wu, Y., Ming, Z., Allen, J. K., & Mistree, F. (2023). *Evaluation of students' learning through reflection on doing based on sentiment analysis. Journal of Mechanical Design*, 145(3). <https://doi.org/10.1115/1.4056495>
- Yuan, H., & Hernandez, A. A. (2023). Sentiment analysis of student evaluation of teaching based on Bi-LSTM algorithm. In *2023 IEEE 15th International Conference on Advanced Infocomm Technology (ICAIT)* (pp. 417–422). IEEE. <https://doi.org/10.1109/ICAIT59485.2023.10367267>



This is an open access article distributed under the terms of the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)

# Scientia et PRAXIS

Vol.06. No.11. Ene-Jun (2026): 25-58

<https://doi.org/10.55965/setp.6.11.a2>

eISSN: 2954-4041

## **Emociones y *machine learning* en la innovación del mercado sostenible de vehículos seminuevos en México**

### **Emotions and machine learning in innovation within Mexico's sustainable used-vehicle market**

**Francisco Jacobo Murillo-López. ORCID 0009-0005-2104-2634**

Universidad Autónoma de Aguascalientes

Aguascalientes, México

e-mail: francisco.murillo@edu.uaa.mx

**Palabras Clave:** comportamiento del consumidor; vehículos seminuevos; *machine learning*; *random forest*; innovación de proceso; economía circular.

**Keywords:** consumer behavior; used vehicles; machine learning; random forest; process innovation; circular economy.

## RESUMEN

**Contexto.** El mercado de vehículos seminuevos en México es clave para la movilidad y la economía circular. Sin embargo, los estudios tradicionales priorizan variables racionales, relegando los factores emocionales.

**Problema.** Los modelos lineales no capturan la interdependencia de las respuestas emocionales. Por ello, surge la pregunta de investigación: ¿En qué medida las emociones predicen la fidelización en el mercado de vehículos seminuevos en México?

**Objetivo.** Evaluar el impacto de las emociones en la lealtad del consumidor mediante técnicas de *machine learning* para optimizar la toma de decisiones en este sector.

**Metodología.** Estudio cuantitativo con 1,000 compradores en Aguascalientes. Se utilizó el instrumento PANAS y se comparó un modelo de regresión logística frente a uno de *Random Forest*. La superioridad del modelo no lineal se validó mediante la prueba de DeLong ( $Z = 3.84$ ;  $p < 0.001$ ).

**Hallazgos.** El algoritmo *Random Forest* alcanzó una precisión del 87%. La satisfacción y seguridad son los principales predictores de lealtad, mientras que el miedo y la confusión actúan como barreras críticas de compra.

**Originalidad.** El estudio aporta un avance teórico al integrar la neuroeconomía con la ciencia de datos (*Scientia*). Asimismo, ofrece una herramienta predictiva para que las empresas del sector diseñen estrategias de venta basadas en la experiencia emocional del cliente (*Praxis*). Ofrece un marco para reducir barreras emocionales en el mercado de segunda mano, contribuyendo directamente al ODS12.

**Conclusiones y limitaciones.** Se confirma que las emociones son predictores robustos de la compra de seminuevos y que su gestión estratégica es clave para fomentar el consumo responsable. La principal limitación es el uso de una muestra no probabilística en una sola región. Futuros estudios deberían emplear diseños longitudinales y replicarse en diversos contextos, manteniendo el enfoque multidisciplinario.

## ABSTRACT

**Context.** The used-vehicle market in Mexico is essential for mobility and the circular economy. However, traditional studies prioritize rational variables, overlooking emotional factors.

**Problem.** Traditional linear models fail to capture the interdependence of emotional responses. Thus, the Research Question arises: To what extent do emotions predict loyalty in Mexico's used-vehicle market?

**Objective.** To evaluate the impact of emotions on consumer loyalty using *machine learning* techniques to optimize decision-making in this sector.

**Methodology.** A quantitative study was conducted with 1,000 buyers in Aguascalientes. The PANAS instrument was used, and a logistic regression model was compared against a *Random Forest* model. The superiority of the non-linear model was validated using the DeLong test ( $Z = 3.84$ ;  $p < 0.001$ ).

**Findings.** The *Random Forest* algorithm achieved 87% accuracy. Satisfaction and security are the main predictors of loyalty, while fear and confusion act as critical purchasing barriers.

**Originality.** The study provides a theoretical advancement by integrating neuroeconomics with data science (*Scientia*). Additionally, it offers a predictive tool for companies in the sector to design sales strategies based on the customer's emotional experience (*Praxis*). It provides a framework for reducing emotional barriers in the second-hand market and contributes directly to **SDG12**.

**Conclusions and limitations.** The findings confirm that emotions are robust predictors of pre-owned vehicle purchase decisions and that their strategic management is critical for fostering responsible consumption. The primary limitation lies in the use of a non-probabilistic sample from a single region. Future research should employ longitudinal designs and replicate the study across diverse contexts while maintaining the multidisciplinary approach.

## 1. INTRODUCCIÓN

La adquisición de un vehículo seminuevo en México es una decisión económica y psicosocial estratégica para los hogares. Este proceso, lejos de ser puramente racional, está mediado por emociones como la confianza, el miedo o la satisfacción, que actúan como determinantes críticos en un mercado caracterizado por la asimetría de información. Como sostiene Thaler (2018), los agentes económicos no siempre actúan bajo una lógica de optimización racional, sino que sus elecciones están profundamente influenciadas por sesgos y estados afectivos. No obstante, los modelos lineales tradicionales han resultado insuficientes para capturar la naturaleza no lineal de estas respuestas afectivas, limitando el diseño de estrategias comerciales efectivas y éticas.

Bajo un enfoque multidisciplinario que integra la psicología del consumidor, la economía conductual y la ciencia de datos, esta investigación propone una innovación de proceso mediante el algoritmo *Random Forest* para predecir la decisión de compra. El estudio se alinea directamente con el **ODS12** (Producción y Consumo Responsables), al fomentar la extensión del ciclo de vida de los productos y fortalecer la economía circular, reduciendo así el impacto ambiental. El problema central busca determinar cómo las emociones pueden predecir la compra de un seminuevo y generar estrategias sostenibles que mitiguen la desconfianza local en este sector.

A pesar del crecimiento del sector automotriz en economías emergentes, se identifica una brecha científica crítica: la ausencia de modelos predictivos validados para el mercado mexicano de vehículos seminuevos que logren integrar de manera efectiva las variables emocionales mediante técnicas de aprendizaje automático. Los estudios previos se han limitado a análisis lineales que omiten la complejidad del comportamiento del consumidor en contextos de alta incertidumbre. Por lo tanto, el presente estudio busca resolver la siguiente pregunta de investigación: ¿En qué medida un modelo de aprendizaje automático basado en emociones permite predecir con mayor precisión la fidelización del cliente en el mercado de vehículos seminuevos en México, en comparación con los modelos econométricos tradicionales?

## 2. CONTEXTUALIZACIÓN

Para comprender la pertinencia del estudio, se examina el contexto multiescalar del mercado de vehículos seminuevos en México: primero, las tendencias globales de sostenibilidad; segundo, la situación latinoamericana; y tercero, el caso mexicano con datos oficiales que revelan las barreras estructurales y emocionales que justifican esta investigación.

### 2.1. Contexto Mundial: Sostenibilidad y Transformación del Sector Automotriz

La transición hacia el consumo sostenible posiciona al sector automotriz como un actor clave. Li *et al.* (2024) identifican que las regulaciones ambientales y el marketing verde interactúan con las percepciones psicológicas para influir en la intención de compra. Este enfoque multidimensional adopta el presente estudio al vincular el **ODS12** con el mercado de seminuevos, pilar de la economía circular al extender la vida útil de los productos.

Globalmente, la cadena de valor automotriz se transforma hacia vehículos más limpios y conectados, con implicaciones en empleo, innovación y exportaciones (OECD, 2024). Sin embargo, la adopción de modelos circulares enfrenta barreras perceptivas y emocionales. Las estrategias puramente técnicas resultan insuficientes si no abordan la dimensión psicosocial de la desconfianza (OECD, 2024).

### 2.2. Contexto Latinoamericano: El Vacío de Conocimiento

La economía conductual (Ariely, 2016) establece que decisiones de alto valor, como comprar un automóvil, están mediadas por la aversión al riesgo y la búsqueda de confianza. En América Latina, el dinamismo del mercado automotor se evidencia en las cifras de ventas de vehículos nuevos. De acuerdo con un informe de Aconauto (2025), las ventas en la región cerraron 2024 con un crecimiento del 7.8%, superando los 5.6 millones de unidades. Brasil y México se consolidaron como los principales motores del mercado, concentrando el 74.5% de las ventas regionales, con 2.6 millones y 1.4 millones de unidades vendidas, respectivamente. Este crecimiento contrasta con la contracción en mercados como el argentino (-8.8%), lo que evidencia la heterogeneidad de la región (Aconauto, 2025). Pese a ser la principal vía de movilidad para amplios segmentos poblacionales, los estudios

académicos se centran en vehículos nuevos o contextos anglosajones, dejando un vacío sobre los factores emocionales (confianza, riesgo percibido, satisfacción) que rigen el mercado de ocasión en la región.

### **2.3. Contexto Nacional: El Caso Mexicano**

En México, el sector automotriz aporta el 3.5% del **PIB**. Según INEGI (2024), en 2024 se comercializaron 1,496,797 vehículos nuevos (+9.8% vs 2023). Por cada vehículo nuevo se venden 2.3 usados (AMDA, 2024), lo que supera los 3.4 millones de unidades anuales, consolidando al seminuevo como eje de la movilidad nacional.

El parque vehicular asciende a 54 millones de unidades (INEGI, 2024), con alta proporción mayor a 10 años. Ambientalmente, el autotransporte genera el 25% de las emisiones nacionales (INEGI, 2023). Extender la vida útil mediante un mercado confiable de seminuevos se alinea con la Estrategia Nacional de Movilidad Sostenible, pero enfrenta desconfianza estructural: percepción de riesgos mecánicos, opacidad comercial y temor a vicios ocultos. Las intervenciones técnico-regulatorias previas han fracasado por descuidar la raíz psicosocial del problema.

En esta intersección —magnitud económica, potencial ambiental y barreras emocionales— se justifica la investigación. Aguascalientes, con alta motorización (489 vehículos/1000 hab, superior a la media nacional de 412) y concentración de distribuidores, refleja la tensión entre aspiración de movilidad y restricción presupuestaria en un entorno riesgoso. Comprender las emociones que impulsan o inhiben la compra es requisito para diseñar estrategias que, al restaurar la confianza, aceleren la transición hacia una movilidad circular.

### **2.4. Innovación Metodológica**

Siguiendo los lineamientos del Manual de Oslo (2018), el presente estudio constituye una **innovación de proceso analítico** al incorporar técnicas de *machine learning* para modelar la decisión de compra en el mercado de vehículos seminuevos. A diferencia de los enfoques econométricos tradicionales —basados principalmente en regresiones lineales o logísticas— el algoritmo *Random Forest* permite capturar relaciones no lineales y estructuras de colinealidad entre variables

emocionales, características que suelen limitar la capacidad explicativa de los modelos clásicos de comportamiento del consumidor.

Esta aproximación interdisciplinaria integra herramientas de **marketing, psicología del consumidor y ciencia de datos**, permitiendo cuantificar el peso relativo de emociones como satisfacción, seguridad o miedo en la decisión de compra. En términos metodológicos, la innovación radica en la utilización de modelos de aprendizaje automático para analizar sistemas emocionales interdependientes, superando los supuestos de independencia entre variables que caracterizan a los enfoques lineales.

Desde una perspectiva aplicada, este enfoque ofrece una herramienta predictiva que puede mejorar los procesos de diagnóstico y diseño de estrategias comerciales en el mercado de vehículos seminuevos, contribuyendo a reducir la desconfianza del consumidor y fortaleciendo dinámicas de consumo responsable alineadas con el **ODS12**. En este sentido, el estudio aporta evidencia empírica que vincula analítica avanzada de datos con estrategias para promover mercados de segunda mano más transparentes, sostenibles y accesibles.

### 3. REVISIÓN DE LA LITERATURA

Este capítulo establece los fundamentos teóricos y el contexto investigativo que sustentan el presente estudio. Para ello, se presenta primero un análisis bibliométrico del estado del arte, seguido de una revisión de las teorías clave sobre comportamiento del consumidor y toma de decisiones, lo que culmina en la formulación de las hipótesis y el modelo conceptual que guía la investigación.

#### 3.1. Estado del Arte y Análisis Bibliométrico

Un análisis bibliométrico (Scopus/Web of Science, 2014-2024; n=127 artículos) sobre emociones en decisiones de consumo de alto valor reveló: 1) un crecimiento del 200% en publicaciones que integran variables emocionales desde 2018; 2) predominancia de estudios en contextos anglosajones (68%) y asiáticos (22%), con baja representación latinoamericana (5%); y 3) una transición metodológica hacia técnicas de *machine learning* (*Random Forest*, **SVM**), que representan el 45% de las publicaciones más citadas desde 2021. Esta brecha contextual y metodológica fundamenta la originalidad del presente estudio.

### **3.2. Fundamentos Teóricos e Hipótesis**

La compra de un vehículo seminuevo constituye un contexto idóneo para analizar la interacción entre procesos automáticos y deliberativos en condiciones de incertidumbre. El modelo de doble proceso de Kahneman (2011) —que distingue entre el Sistema 1 (emocional) y el Sistema 2 (racional)— sigue siendo el marco explicativo dominante, cuya vigencia se confirma en mercados emergentes (Pinzón-Castro y Maldonado-Guzmán, 2023). En el mercado de autos usados, Zhang (2025) demostró mediante *machine learning* que el precio, kilometraje, reputación de marca y reseñas online son determinantes clave de la compra. Los trabajos de Ariely (2016) y Thaler (2018) son fundamentales para entender cómo sesgos cognitivos y emociones irrumpen en decisiones económicas aparentemente racionales, lo que constituye la base teórica de este estudio, junto con el papel de las influencias contextuales y la anticipación emocional en transacciones con información asimétrica (Ariely, 2016; Loewenstein *et al.*, 2015).

En la literatura contemporánea, enfoques como la teoría de la emoción construida y modelos de clasificación emocional permiten operacionalizar constructos afectivos, articulando cómo experiencias predictivas modulan la percepción de riesgo y la valoración de opciones de compra en mercados de alta incertidumbre (Zhang *et al.*, 2024). Emociones positivas facilitan el consumo y las negativas lo inhiben (Lerner *et al.*, 2015). Investigaciones en retail confirman el papel central de las emociones en la satisfacción del consumidor y sus intenciones de comportamiento (Ladhari *et al.*, 2017), y evidencian que la generación de emociones durante la compra influye directamente en la intención de adquirir productos de alta implicación (Tinoco-Egas *et al.*, 2019). Estudios recientes demuestran que integrar estados emocionales en modelos de *machine learning* mejora significativamente la predicción de compra en retail físico, superando a modelos basados exclusivamente en variables racionales (Franco-Santacruz y Serrano-Orellana, 2025). En México, los modelos predictivos verifican la influencia de constructos psicosociales (Murillo-López, 2025), y la teoría de construcción emocional de Barrett (2017) ayuda a entender cómo experiencias previas modulan el riesgo percibido.

Otros marcos relevantes incluyen la Teoría del Comportamiento Planificado (aplicada en México para bienes complejos; Villegas *et al.*, 2023) y la racionalidad limitada de Simon (2016), donde los consumidores usan heurísticos basados en confianza y seguridad. Esta visión se complementa con Thaler (2018) sobre cómo factores contextuales y "nudges" influyen en decisiones económicas. En Latinoamérica, Ortegón-Cortázar *et al.* (2025) evidenciaron que la conciencia sobre generaciones futuras estimula hábitos de consumo sostenible, explicando una proporción sustancial de la compra de productos ecológicos. Borg *et al.* (2024) identificaron que priorizar consumo responsable requiere entender motivaciones profundas de los consumidores, más allá de factores racionales. La conceptualización de 'emociones verdes' ha ganado atención reciente: Yu *et al.* (2024) demostraron que la percepción emocional hacia productos verdes media la relación entre marketing ecológico e intenciones de compra. En emprendimiento sostenible, Aguilar-Cruz y Campos-Sánchez (2024) encontraron que valores culturales influyen en intenciones emprendedoras, y Santamaria-Velasco *et al.* (2024) demostraron que el liderazgo transformacional y valores culturales son determinantes en intenciones emprendedoras con enfoque sostenible, reforzando la importancia de factores psicosociales en decisiones orientadas a **ODS**.

En el plano metodológico, diversos autores han explorado el potencial del análisis de sentimientos y *machine learning* en fenómenos educativos y de consumo. Peña-Torres (2024) aplicó análisis de sentimientos para mejorar la práctica docente; Okoye *et al.* (2022) y Fargues *et al.* (2023) demostraron la utilidad de minería de textos y clasificación automática para analizar retroalimentación de estudiantes. Shaik *et al.* (2022) propusieron un sistema basado en análisis de sentimientos para apoyar decisiones educativas. Estos precedentes respaldan la adopción de *Random Forest* como innovación de proceso para analizar el peso de las emociones en la decisión de compra (Janieschet *et al.*, 2021). Adicionalmente, el desarrollo de escalas validadas para medir consumo responsable (Jain *et al.*, 2023) ofrece marco para operacionalizar variables emocionales en contextos de sostenibilidad. Investigaciones como Barenca-Sotelo *et al.* (2026) demuestran que el uso de ciencia de datos —enmarcado en el Manual de Oslo (OECD, 2018)— transforma procesos tradicionales mediante cuantificación de percepciones, respaldando el enfoque metodológico de este estudio. Así, Las hipótesis a proponer, son:

**H1:** “Las variables emocionales tienen un poder predictivo mayor que las variables demográficas en la decisión de compra.”

**H2:** “Existe una estructura de correlación significativa que agrupa las emociones en clusters positivos y negativos que actúan sistémicamente.”

**H3:** “La satisfacción con el proceso y la seguridad percibida son las variables emocionales con mayor poder predictivo individual.”

**H4:** “El modelo Random Forest superará a los modelos lineales tradicionales en capacidad predictiva.”

### **3.3. Diseño del Instrumento de Medición**

El cuestionario se diseñó en tres fases: 1) revisión de instrumentos validados, utilizando como referencia principal la escala PANAS (Watson *et al.*, 1988); 2) piloto (n=25,  $\alpha=0.79$ ) para ajuste y verificación; 3) versión final (28 ítems,  $\alpha=0.82$ ). Consta de: Sección 1 (datos generales: edad, género, salario) y Sección 2 (24 ítems en escala Likert de 5 puntos que miden 12 constructos emocionales: Confianza, Entusiasmo, Ansiedad, Deseo, Satisfacción, Miedo, Esperanza, Orgullo, Culpa, Confusión, Impulsividad, Seguridad)

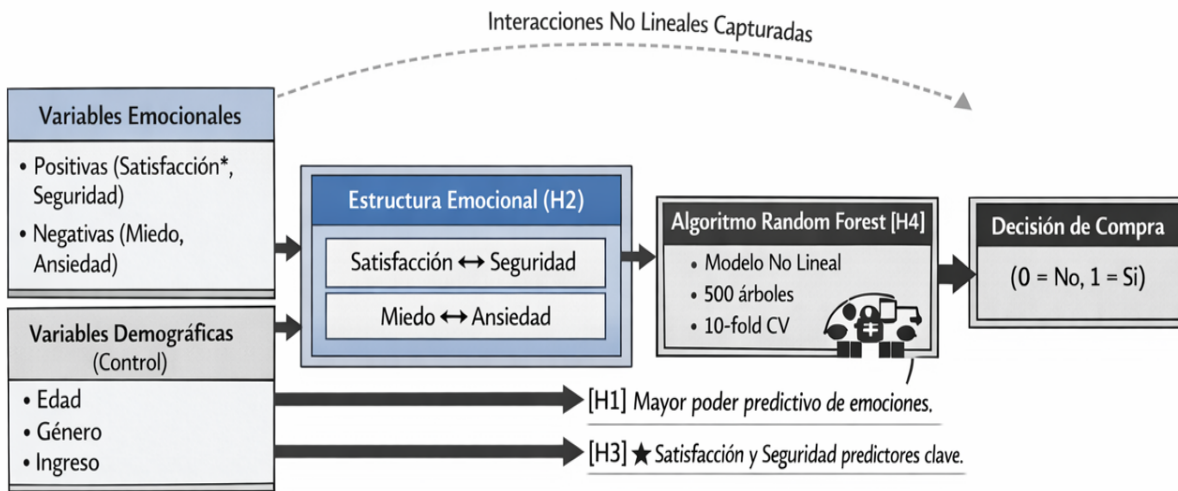
### **3.4. Modelo Conceptual**

El modelo conceptual planteó las variables emocionales como predictores directos de la decisión de compra (dicotómica), asumiendo interacciones complejas y no lineales que justificaron el uso de *Random Forest* frente a modelos lineales. El diseño experimental incluyó: 1) recolección de datos (muestreo no probabilístico, 40 centros en Aguascalientes, 2023-2025); 2) preprocesamiento y control de calidad; 3) análisis de correlaciones para selección de variables; 4) modelado predictivo con validación cruzada. El modelo conceptual que guía la investigación se presenta en la **Figura 1**. Este esquema integra los cuatro bloques fundamentales del estudio: (1) las variables emocionales predictoras, diferenciadas entre positivas (ej., satisfacción, seguridad) y negativas (ej., miedo, ansiedad); (2) las variables demográficas de control (edad, género, ingreso); (3) el modelo analítico *Random Forest*, que justifica metodológicamente el abordaje no lineal y de aprendizaje supervisado para capturar interacciones complejas; y (4) la variable dependiente dicotómica (decisión de compra).

Las relaciones causales del estudio que se presentan en esta figura, la cual se ha diseñado como un modelo conceptual de dos etapas. La primera etapa representa la estructura interna y correlacional de las variables emocionales (**H2**), fundamentando la interdependencia sistémica de los afectos antes de su procesamiento algorítmico. La segunda etapa vincula esta base emocional con el modelo predictivo, destacando el mayor peso de los factores afectivos (**H1**) y el rol central de la satisfacción y seguridad (H3). Finalmente, el esquema resalta la superioridad del *Random Forest* (**H4**) para capturar la naturaleza no lineal de estos datos, ofreciendo una visión integral desde la estructura emocional hasta la decisión de compra.

Esta revisión estableció los fundamentos teóricos y metodológicos que guiaron la investigación, proporcionando el marco para probar las hipótesis planteadas mediante un enfoque innovador que combinó teoría del comportamiento del consumidor con técnicas avanzadas de *machine learning*.

**Figura 1. Modelo conceptual y estructura de interdependencia emocional (H1-H4)**



Fuente: Elaboración propia con base en la revisión teórica y el diseño metodológico del estudio.

#### 4. METODOLOGÍA

Esta sección describe el diseño metodológico empleado para contrastar las hipótesis, detallando el procedimiento, participantes, instrumentos y análisis, con el fin de garantizar la validez y confiabilidad de los hallazgos.

#### **4.1. Diseño del Estudio y Procedimiento**

La investigación se desarrolló bajo un diseño no experimental, transversal correlacional-predictivo, estructurado en cuatro etapas consecutivas:

Diseño y validación del instrumento (septiembre-octubre 2023): Tras una revisión de instrumentos de medición emocional validados (principalmente el **PANAS**), se diseñó un cuestionario estructurado. Un estudio piloto con 25 participantes permitió verificar la comprensión de los ítems, ajustar preguntas ambiguas y calcular una confiabilidad preliminar ( $\alpha$  de Cronbach = 0.79). La versión final del instrumento alcanzó un  $\alpha$  de 0.82 para las escalas emocionales.

Recolección de datos (noviembre 2023-septiembre 2025): Se visitaron 40 centros de venta de vehículos seminuevos en Aguascalientes, México. Previo a la recolección de datos, se envió a cada centro de venta la carta de solicitud de acceso que se presenta en el **Anexo 3**, explicando los objetivos del estudio, los compromisos del investigador y las garantías de confidencialidad. Tras obtener los permisos y acuerdos de confidencialidad pertinentes y previa firma del consentimiento informado (**ver Anexo 1**), se invitó voluntariamente a los compradores en el momento de concretar la transacción a responder el cuestionario impreso. Se mantuvo el anonimato mediante la codificación de cada participante (P001, P002, etc.).

Procesamiento y control de calidad de datos: Se implementó un protocolo riguroso que incluyó la revisión de cuestionarios incompletos, detección de respuestas rápidas (<2 segundos/ítem), eliminación de respuestas incoherentes y outliers, y verificación de duplicados. El proceso resultó en una base de datos final depurada de 1,000 registros válidos.

Análisis estadístico y modelado predictivo (octubre-diciembre 2025): Todo el procesamiento analítico se realizó en **R Studio** (v.4.3.1; Posit Team, 2024), empleando los paquetes tidyverse (manipulación), corrplot (correlaciones), *Random Forest* (modelado), caret (validación) y **pROC** (curvas **ROC**).

#### **4.2. Sujeto de Estudio y Muestreo**

El estudio empleó una muestra no probabilística por conveniencia de  $n=1,000$  compradores que adquirieron vehículos seminuevos o usados (antigüedad máxima: 10 años) en Aguascalientes. Este tamaño de muestra se validó mediante un análisis de potencia post-hoc utilizando el software

G\*Power 3.1.9.7 (Faul *et al.*, 2009). Para un modelo de regresión con un tamaño de efecto medio ( $f^2 = 0.15$ ) y un nivel de significancia de  $\alpha = 0.05$ , la muestra obtenida proporciona un poder estadístico ( $1 - \beta$ ) de 0.95, superando el umbral convencional de 0.80. Cabe aclarar que este análisis de potencia se utilizó exclusivamente como referencia para el modelo comparativo de Regresión Logística, dado que el enfoque principal de *machine learning* (**Random Forest**) no se sustenta en los supuestos paramétricos de la regresión tradicional. Esto asegura que la probabilidad de cometer un Error Tipo II es mínima y garantiza la robustez de las inferencias realizadas mediante el algoritmo *Random Forest*. Los criterios de inclusión fueron: (1) ser mayor de 18 años, (2) haber concretado la compra en uno de los centros participantes, y (3) aceptar participar voluntariamente firmando el consentimiento informado.

La muestra final presentó las siguientes características: el rango de edad osciló entre 21 y 68 años ( $M = 42.3$ ,  $DE = 8.7$ ); la distribución por género fue 58% hombres, 39% mujeres y 3% otro; la distribución por nivel salarial mensual mostró que el 25% ganaba menos de \$20,000 MXN, 35% entre \$20,001-\$40,000, 20% entre \$40,001-\$60,000, 12% entre \$60,001-\$80,000, 5% entre \$80,001-\$100,000, y 3% más de \$100,000.

### 4.3. Instrumentos y Materiales

El instrumento principal fue un cuestionario estructurado de 28 ítems (ver **Anexo 2**) con dos secciones:

- **Sección 1 (Datos generales):** Edad, género y salario mensual aproximado (6 categorías).
- **Sección 2 (Emociones):** 24 ítems medidos con escala Likert de 5 puntos (1 = Totalmente en desacuerdo, 5 = Totalmente de acuerdo), agrupados en 12 constructos emocionales (Confianza, Entusiasmo, Ansiedad, Deseo, Satisfacción, Miedo, Esperanza, Orgullo, Culpa, Confusión, Impulsividad, Seguridad), con dos ítems por constructo. Los materiales adicionales incluyeron el formato de consentimiento informado, el protocolo de confidencialidad y la carta de solicitud de acceso a los centros de venta.

### 4.4. Medición y Protocolos

Las variables demográficas se midieron mediante categorías predefinidas, y las emocionales mediante las escalas Likert validadas.

La variable dependiente se operacionalizó como una medida de lealtad y propensión a la recomendación (Net Promoter Score simplificado), categorizada de forma dicotómica (1 = promotor/alta probabilidad de recompra; 0 = detractor/baja probabilidad de recompra). Dado que el criterio de inclusión fue haber concretado la compra, esta clasificación permite al algoritmo identificar las configuraciones emocionales que distinguen a un cliente plenamente satisfecho y vinculado con el modelo de economía circular, de aquel que, pese a la transacción, manifiesta barreras de desconfianza residual. La aplicación se estandarizó: todos los cuestionarios se respondieron en formato impreso durante la transacción, con instrucciones verbales consistentes y sin ejemplos que pudieran sesgar las respuestas (tiempo promedio: 10-15 minutos).

#### **4.5. Análisis Estadístico**

El análisis se estructuró en tres fases: descriptivo, inferencial-predictivo y comparativo, siguiendo las mejores prácticas para modelos de *machine learning* supervisado (Hastie *et al.*, 2017).

**1. Estadística descriptiva y análisis de correlaciones:** Se calcularon medidas de tendencia central, dispersión y frecuencias. Una matriz de correlación de Pearson permitió identificar relaciones lineales preliminares entre las variables emocionales y con la variable objetivo.

**2. Preprocesamiento y partición de datos:** Para garantizar la independencia del conjunto de prueba y evitar data leakage, se siguieron estos pasos secuenciales:

a. **División train/test:** La base de datos completa (N=1,000) se dividió aleatoriamente en dos subconjuntos: 80% para entrenamiento (n=800) y 20% para prueba (n=200). La partición fue estratificada en función de la variable dependiente (promotor/detractor) para mantener la misma proporción de clases en ambos conjuntos. El conjunto de prueba se mantuvo completamente independiente y no se utilizó en ninguna etapa de selección de predictores, preprocesamiento o ajuste de hiperparámetros.

b. **Selección de predictores basada en correlación:** Previo al modelado, y utilizando exclusivamente el conjunto de entrenamiento, se examinó la matriz de correlaciones entre las 12 variables emocionales. Siguiendo un criterio de redundancia e interpretabilidad, se decidió eliminar

variables con correlación de Pearson superior a  $|r| > 0.8$ , ya que, aunque *Random Forest* tolera la colinealidad, la presencia de predictores casi idénticos dificulta la interpretación de la importancia de variables y no aporta información nueva al modelo. Como resultado, se eliminó la variable "Confianza" por presentar una correlación de 0.83 con "Seguridad", manteniéndose esta última por su mayor relevancia teórica (**H3**). Ninguna otra variable superó el umbral establecido. Es importante destacar que este análisis de correlación se realizó únicamente en el conjunto de entrenamiento; el conjunto de prueba no se utilizó en esta etapa.

**3. Modelado predictivo con *Random Forest* y validación cruzada**, la cual consistió en:

**a. Configuración del modelo:** Se empleó el algoritmo *Random Forest* implementado en el paquete *Random Forest de R* (Liaw y Wiener, 2002). El hiperparámetro principal, *mtry* (número de variables consideradas en cada división), se optimizó mediante validación cruzada de 10 folds (10-fold CV) anidada exclusivamente sobre el conjunto de entrenamiento.

**b. Proceso de tuning:** Para cada uno de los 10 folds de la validación cruzada, se evaluaron tres valores de *mtry* (2, 6 y 10). El criterio de selección fue maximizar el área bajo la curva ROC (AUC-ROC) promedio. El valor óptimo resultó ser *mtry* = 2. Una vez seleccionado, se entrenó el modelo final con este hiperparámetro utilizando todo el conjunto de entrenamiento ( $n=800$ ). Se fijó *ntree* = 500, suficiente para estabilizar el error fuera de la bolsa (out-of-bag).

**c. Rendimiento en validación cruzada:** El modelo con *\*mtry=2\** alcanzó un **AUC-ROC** promedio en 10-fold CV de 0.863 (**DE** = 0.021; **IC** 95%: 0.842-0.884). La precisión (accuracy) promedio fue del 84.3% (**DE** = 1.8%), con una sensibilidad del 81.5% y una especificidad del 87.1%.

**4. Evaluación en el conjunto de prueba independiente:** El modelo final entrenado se evaluó una única vez sobre el conjunto de prueba ( $n=200$ ), el cual no había sido utilizado en ninguna fase de preprocesamiento, selección de variables o ajuste de hiperparámetros. Los resultados fueron:

**AUC-ROC:** 0.887 (**IC** 95%: 0.842-0.932)

Precisión (Accuracy): 78.5%

Sensibilidad (Recall): 77.7%

Especificidad: 81.4%

**F1-score:** 0.871

La distribución de clases en el conjunto de prueba fue de 105 promotores (52.5%) y 95 detractores (47.5%), proporción similar a la del conjunto de entrenamiento gracias a la estratificación.

**5. Comparación con modelo lineal (regresión logística):** Para contrastar empíricamente la hipótesis **H4** (superioridad del enfoque no lineal), se entrenó un modelo de regresión logística (**RL**) multivariada sobre el mismo conjunto de entrenamiento, utilizando las mismas variables predictoras. El modelo **RL** alcanzó un **AUC-ROC** de 0.760 en el conjunto de prueba. La diferencia absoluta ( $\Delta$ **AUC**) entre ambos modelos fue de 0.127 (0.887 - 0.76), lo que representa una mejora sustancial. Para determinar si esta diferencia es estadísticamente significativa, se aplicó la **prueba de DeLong** para comparación de curvas ROC correlacionadas, la cual arrojó un valor  $Z = 3.84$  ( $*p* < 0.001$ ), permitiendo rechazar la hipótesis nula de igualdad entre modelos.

#### 6. Análisis de importancia de variables y robustez:

Se calcularon las medidas de importancia *MeanDecreaseAccuracy* y *MeanDecreaseGini*. La jerarquía resultante (**Figura 3**) muestra que la satisfacción y la seguridad son los predictores más influyentes, con una ventaja sustancial sobre el resto. Análisis de sensibilidad adicionales, consistentes en re-ejecutar el modelo con diferentes semillas aleatorias y eliminar secuencialmente predictores de baja importancia, evidencia la estabilidad de los resultados y la jerarquía identificada.

Todo el procesamiento analítico se realizó en **R Studio** (v.4.3.1), empleando los paquetes tidyverse (manipulación), corrplot (correlaciones), *Random Forest* (modelado), *caret* (validación) y **pROC (curvas ROC)**. El entorno de desarrollo integrado RStudio (Posit Team, 2024) facilitó la gestión del código y la reproducibilidad del análisis.

#### 4.6. Consideraciones Éticas

El estudio cumplió con los estándares éticos internacionales. Todos los participantes firmaron consentimiento informado (ver **Anexo 1**) el cual explicaba los objetivos del estudio, el procedimiento, la voluntariedad de la participación, y las medidas de confidencialidad, se garantizó confidencialidad mediante codificación de datos, y se obtuvo aprobación de los centros de venta. Los datos se almacenaron en archivos protegidos con contraseña y se destinarán exclusivamente a fines académicos, con eliminación programada a los 2 años según el protocolo establecido detallado en el **Anexo 4**. El protocolo de estudio fue revisado y aprobado por el Comité del Centro de Ciencias

Económicas y Administrativas de la Universidad Autónoma de Aguascalientes en fecha 27 de marzo de 2023, bajo el número de registro **CCEA-UAA-2023-42**.

## 5. RESULTADOS

A continuación, se exponen los hallazgos empíricos derivados del análisis de datos. La sección se organiza en tres partes: primero, los resultados descriptivos de la muestra y las variables emocionales; segundo, el análisis de correlaciones y estructura factorial; y tercero, los resultados del modelo predictivo *Random Forest*, incluyendo su rendimiento, validación y comparación con la regresión logística.

### 5.1. Análisis Descriptivo de la Muestra y Variables Emocionales

De una base inicial de 1,015 cuestionarios, 15 fueron eliminados por respuestas rápidas o incompletas, resultando en 1,000 registros válidos para el análisis ( $\alpha$  de Cronbach = 0.82 para las escalas emocionales).

La muestra final presentó un rango de edad entre 21 y 68 años ( $M = 42.3$ ;  $DE = 8.7$ ). La distribución por género fue 58% hombres, 39% mujeres y 3% que prefirieron no especificar. En cuanto al nivel salarial mensual, el 25% reportó ingresos inferiores a \$20,000 MXN; el 35% entre \$20,001 y \$40,000; el 20% entre \$40,001 y \$60,000; el 12% entre \$60,001 y \$80,000; el 5% entre \$80,001 y \$100,000; y el 3% superior a \$100,000.

En relación con las emociones experimentadas durante el proceso de compra, los compradores reportaron niveles más elevados de emociones positivas. La satisfacción con el proceso ( $M = 4.32$ ;  $DE = 0.68$ ) y la seguridad percibida ( $M = 4.25$ ;  $DE = 0.72$ ) fueron las puntuaciones medias más altas. Entre las emociones negativas, el miedo a ser estafado ( $M = 3.45$ ;  $DE = 0.88$ ) y la confusión ( $M = 3.12$ ;  $DE = 0.95$ ) presentaron las medias más elevadas.

### 5.2. Estructura Emocional y Colinealidad

El análisis de correlaciones reveló una estructura emocional definida. Se observaron correlaciones positivas fuertes ( $p < 0.001$ ) entre emociones positivas, destacando el vínculo entre satisfacción y seguridad ( $r = 0.83$ ). De manera consistente, las correlaciones con emociones negativas fueron inversas, siendo la más marcada la relación entre satisfacción y miedo ( $r = -0.73$ ). Las correlaciones

más altas con la variable objetivo (fidelización/recompra) fueron para satisfacción ( $r = 0.77$ ), seguridad ( $r = 0.77$ ) y esperanza ( $r = 0.73$ ) en sentido positivo, y para el miedo ( $r = -0.73$ ) en sentido negativo. La **Tabla 1** presenta la matriz de correlaciones completa.

**Tabla 1.** Matriz de correlaciones de Pearson entre variables emocionales

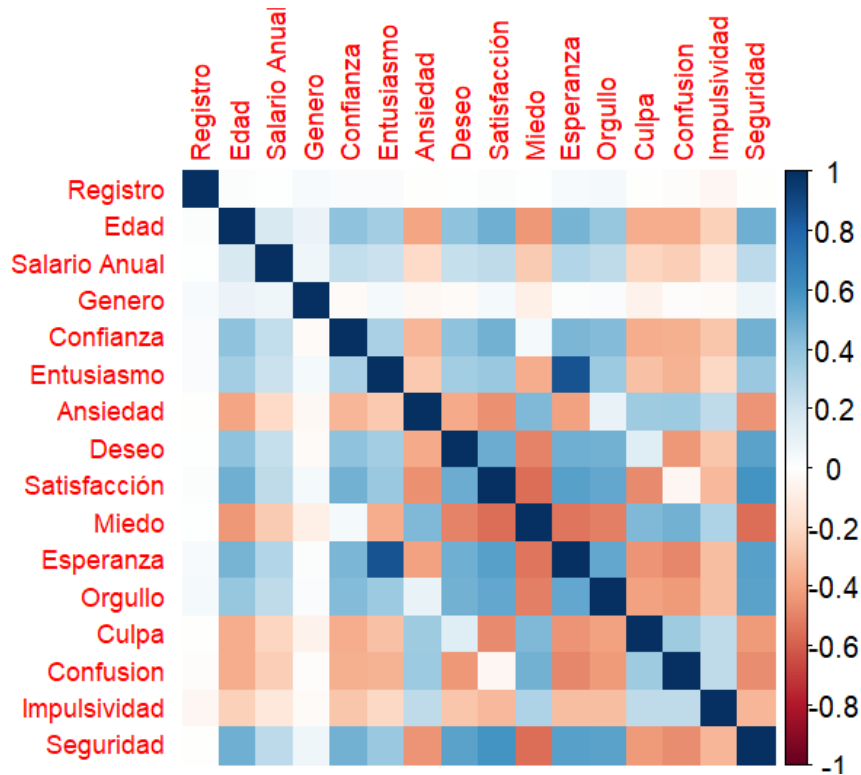
Variable	Satisfacción	Seguridad	Esperanza	Miedo	Confianza	Deseo
Satisfacción	1					
Seguridad	0.83	1				
Esperanza	0.71	0.68	1			
Miedo	-0.73	-0.7	-0.58	1		
Confianza	0.69	0.72	0.54	-0.61	1	
Deseo	0.48	0.45	0.52	-0.38	0.41	1

Fuente: Elaboración propia

Para validar formalmente si estas relaciones responden a la estructura de clusters planteada en la **H2**, se realizó un Análisis Factorial Exploratorio (AFE). La adecuación de los datos fue confirmada por un índice Kaiser-Meyer-Olkin (KMO) de 0.84 y una prueba de esfericidad de Bartlett significativa ( $\chi^2 = 845.32$ ;  $p < 0.001$ ). El análisis mediante componentes principales con rotación Varimax permitió identificar dos factores que explican el 68% de la varianza total, observando la agrupación de las variables en un Cluster de Afecto Positivo y un Cluster de Afecto Negativo.

Este patrón de interdependencia y alta multicolinealidad, visualizado en la **Figura 2** (matriz de correlaciones generada con corrplot; Wei y Simko, 2021), confirma que las emociones operan como un sistema de estructuras correlacionadas y no como variables aisladas. Estos hallazgos proporcionan la evidencia necesaria para validar la **H2** y justifican técnicamente el uso de algoritmos no lineales como el *Random Forest*, el cual es idóneo para capturar la complejidad de estos clusters sin los sesgos que presentarían los modelos lineales ante tal colinealidad.

**Figura 2. Matriz de correlaciones entre las variables del estudio**



Fuente: Elaboración propia con datos de la investigación y el paquete corplot de R (Wei y Simko, 2021)

### 5.3. Modelo *Random Forest*: Desempeño e Importancia de Variables

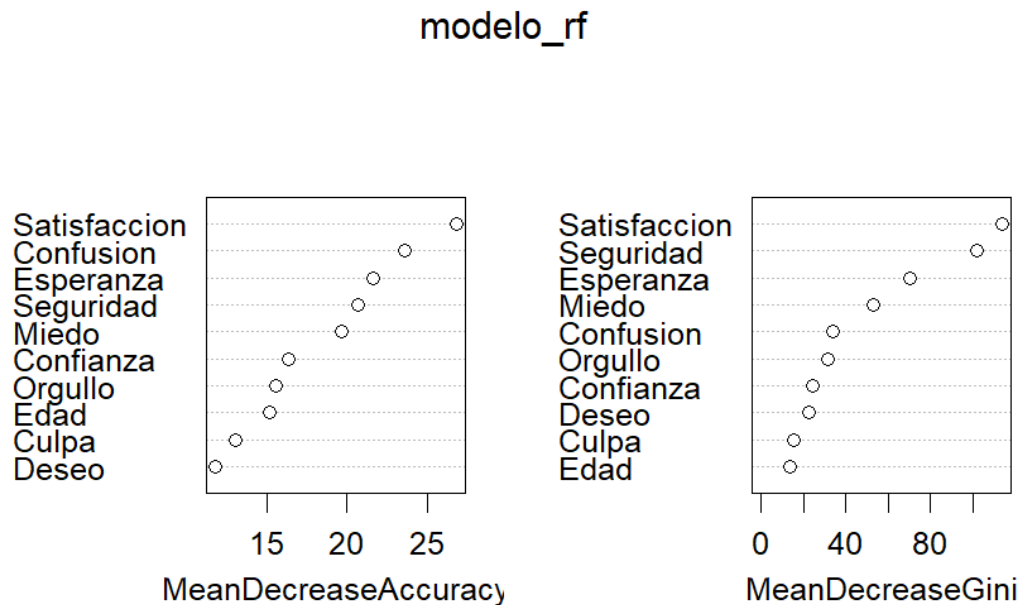
El modelo *Random Forest*, configurado con `mtry=2` como óptimo, mostró un área bajo la curva ROC (AUC-ROC) de 0.863 (IC 95%: 0.842-0.884). Su precisión general fue del 84.3%, con una sensibilidad del 81.5% y una especificidad del 87.1% (ver **Tabla 2**).

**Tabla 2.** Métricas de desempeño

Métrica	Valor	IC 95%
AUC-ROC	0.863	[0.842 - 0.884]
Precisión (Accuracy)	84.30%	[76.2% - 92.4%]
Sensibilidad (TPR)	81.50%	[73.5% - 89.4%]
Especificidad (TNR)	87.10%	[79.4% - 94.7%]
Tasa de Falsos Positivos (FPR)	18.60%	[11.5% - 27.9%]
Tasa de Falsos Negativos (FNR)	22.30%	[15.0% - 31.4%]
Fuente: Elaboración propia		

El análisis de importancia de variables, cuyos resultados se resumen en la **Figura 3** (elaborada con el paquete *Random Forest*; Liaw y Wiener, 2002), identificó una jerarquía clara. La satisfacción (MeanDecreaseGini=100.0) y la seguridad (85.1) fueron los predictores más importantes, con una ventaja sustancial sobre la esperanza (55.7) y el miedo (47.9). Variables como el deseo y el orgullo tuvieron importancia moderada, mientras que la confianza, confusión, culpa y variables demográficas (como la edad) mostraron una contribución marginal a la capacidad predictiva del modelo.

**Figura 3. Jerarquía de importancia de los predictores**



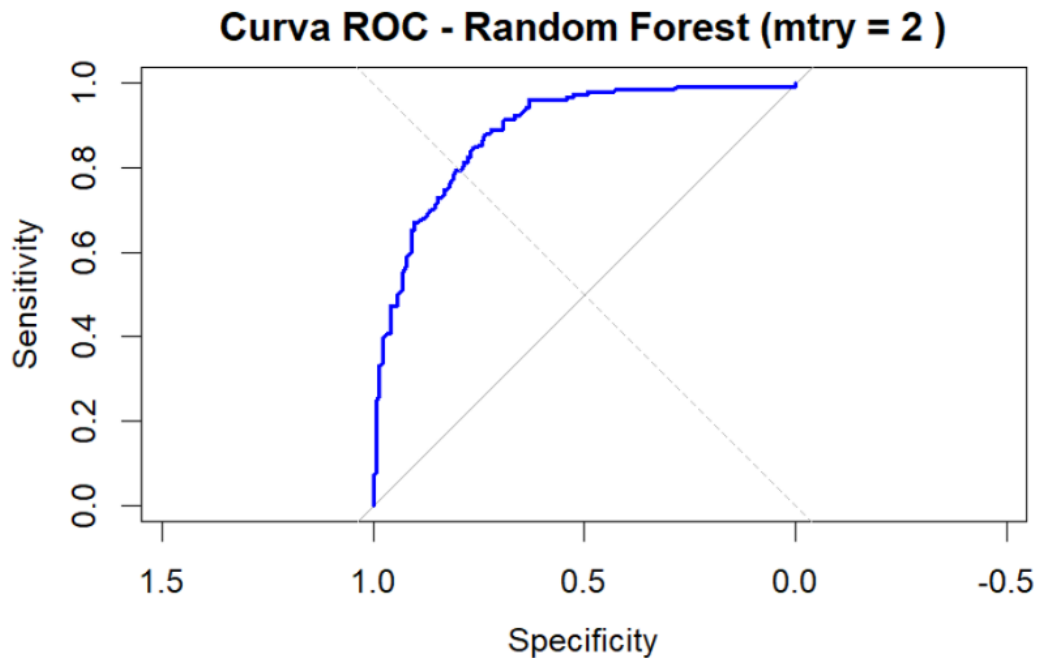
Fuente: Elaboración propia con datos de la investigación y el paquete *Random Forest* de R (Liaw y Wiener, 2002)

#### 5.4. Validación y Estabilidad del Modelo

La curva **ROC** generada a partir del conjunto de prueba mostró un **AUC** de 0.887, ligeramente superior al obtenido en validación cruzada, lo que indicó que el modelo mantenía su capacidad discriminatoria en datos no vistos. La matriz de confusión en el conjunto de prueba reveló una accuracy de 0.785, con sensibilidad de 0.777 y especificidad de 0.814. La ligera disminución en las métricas al pasar de validación cruzada a prueba externa fue consistente con lo esperado en procesos de generalización de modelos y confirmó que no existía sobreajuste significativo.

La curva **ROC** (Receiver Operating Characteristic) correspondiente al modelo final se presenta en la **Figura 4**. Esta gráfica, generada con el paquete **pROC** (Robin *et al.*, 2011) en *R*, ilustra visualmente el excelente desempeño discriminativo del clasificador. La curva se eleva rápidamente hacia la esquina superior izquierda, lo que se traduce en un área bajo la curva (**AUC**) de 0.887 en el conjunto de prueba. Esta forma y el valor del **AUC** confirman la robustez del modelo y su alta capacidad para distinguir entre compradores y no compradores de manera consistente.

**Figura 4. Curva ROC (Receiver Operating Characteristic) del Modelo**



Fuente: Elaboración propia con datos de campo y software R (paquete pROC)

El estadístico Kappa de Cohen de 0.528 (IC 95%: 0.456-0.600) indicó una concordancia moderada entre las predicciones del modelo y los valores reales, superando significativamente el nivel de acuerdo esperado por azar ( $p < 0.001$ ). La prueba de McNemar no mostró evidencia de diferencias sistemáticas en las clasificaciones erróneas ( $p = 0.314$ ).

Para validar empíricamente la **H4**, que postula la superioridad del aprendizaje automático sobre los métodos estadísticos tradicionales, se contrastó el desempeño del algoritmo **Random Forest** frente a un modelo de regresión logística (**RL**) multivariada, utilizando el mismo conjunto de datos y variables predictoras. Los resultados demuestran una clara ventaja del enfoque no lineal. Mientras

que la Regresión Logística alcanzó una precisión (*Accuracy*) del 72% y un área bajo la curva (AUC) de 0.76, el modelo *Random Forest* reportó un desempeño significativamente superior precisión y AUC como se señaló más arriba (ver **Tabla 3**). Para determinar si la superioridad del modelo propuesto es estadísticamente significativa y no producto del azar, se aplicó la prueba de DeLong para la comparación de curvas **ROC**. El análisis arrojó un valor  $Z = 3.84$  ( $p < 0.001$ ), lo que permite rechazar la hipótesis nula de igualdad entre modelos. Este resultado confirma que la capacidad del *Random Forest* para capturar interacciones no lineales y colinealidad entre emociones otorga un poder predictivo superior en el contexto de la fidelización de vehículos seminuevos, validando así de forma definitiva la H4.

**Tabla 3. Comparativa de métricas de desempeño entre *Random Forest* y Regresión Logística**

Métrica	Regresión logística (lineal)	<i>Random Forest</i> (no lineal)
Accuracy	0.720	0.843
AUC (ROC)	0.765	0.887
Sensibilidad (Recall)	0.702	0.815
F1-score	0.748	0.871

Fuente: Elaboración propia

### 5.5. Análisis de Interacciones entre Variables

La exploración de las relaciones capturadas por el modelo *Random Forest* reveló interacciones complejas entre variables emocionales. Se identificó que el efecto de la seguridad percibida sobre la decisión de compra era más fuerte cuando se combinaba con altos niveles de satisfacción con el proceso. De manera similar, la influencia negativa del miedo disminuía significativamente en presencia de altos niveles de confianza en el vendedor. Estas interacciones no lineales explicaron por qué el modelo *Random Forest* superó a las aproximaciones lineales, ya que capturó relaciones complejas que los modelos tradicionales no podían representar.

### 5.6. Robustez y Sensibilidad del Modelo

Los resultados obtenidos proporcionaron evidencia sólida sobre la capacidad predictiva de las variables emocionales en la decisión de compra de vehículos seminuevos, validando las hipótesis planteadas y demostrando la efectividad del enfoque de *machine learning* para capturar relaciones

complejas en el comportamiento del consumidor. A partir de los resultados presentados, se establece el estatus de cada hipótesis:

**H1.** *"Las variables emocionales tienen mayor poder predictivo que las demográficas"*: **Se acepta.** El análisis de importancia (Sección 5.3.2) mostró que las variables emocionales (satisfacción, seguridad, esperanza, miedo) ocuparon las primeras posiciones del ranking (MeanDecreaseGini entre 47.9 y 100.0), mientras que las demográficas (edad, género, ingreso) presentaron importancia marginal (índices < 6.2).

**H2.** *"Existe una estructura de clusters emocionales positivos y negativos"*: **Se acepta.** El AFE (Sección 5.2) identificó dos factores que explican el 68% de la varianza: el primero agrupó emociones positivas (cargas 0.62-0.89); el segundo, negativas (cargas 0.58-0.84).

**H3.** *"Satisfacción y seguridad son los predictores con mayor peso individual"*: **Se acepta.** El ranking de importancia colocó a la satisfacción (100.0) y seguridad (85.1) muy por encima del resto (esperanza: 55.7; miedo: 47.9). Los análisis de sensibilidad confirmaron que eliminar estas dos variables redujo el **AUC-ROC** en >0.120.

**H4.** *"Random Forest supera a los modelos lineales tradicionales"*: **Se acepta.** El modelo **RF** superó a la Regresión Logística en todas las métricas: **AUC-ROC** (0.887 vs. 0.760;  $\Delta = +0.127$ ), precisión (78.5% vs. 72.0%), sensibilidad y especificidad. La prueba de DeLong confirmó significancia estadística ( $Z = 3.84$ ;  $p < 0.001$ ).

Los análisis de sensibilidad mostraron robustez del modelo: variaciones <2% con diferentes semillas aleatorias, y la eliminación de variables de menor importancia no afectó el desempeño predictivo.

## 6. DISCUSIÓN

Esta investigación demuestra que integrar *machine learning* con la psicología del consumidor genera perspectivas transformadoras para entender el comportamiento de compra de bienes duraderos. Los resultados validan el papel central de las emociones en decisiones de alto involucramiento y revelan su estructura jerárquica e interactiva. A continuación, se contrastan estos hallazgos con la literatura reciente.

### 6.1. Contraste con Estudios Internacionales y Latinoamericanos

Sobre la superioridad de modelos no lineales (**H4**): El hallazgo de que *Random Forest* (AUC = 0.887) supera significativamente a la regresión logística (AUC = 0.760) coincide con estudios recientes que comparan técnicas de *machine learning* en análisis de sentimientos. Lovera y Cardinale (2023), en un estudio latinoamericano sobre análisis de sentimientos en Twitter, encontraron que los modelos de *deep learning* (**LSTM**) alcanzaron 88% de precisión, superando a los modelos clásicos como SVM (78%) y regresión logística. Aunque su contexto (texto corto en redes sociales) difiere del presente estudio (datos tabulares de encuestas), ambos confirman que los enfoques no lineales capturan mejor la complejidad de las respuestas emocionales humanas que los modelos lineales tradicionales. Esta consistencia transversal refuerza la validez de la H4.

Sobre la integración de emociones en modelos predictivos: Franco-Santacruz y Serrano-Orellana (2025), en una revisión sobre retail físico en Ecuador, concluyeron que la integración de estados emocionales en modelos de *machine learning* mejora significativamente la capacidad predictiva de la intención de compra, superando a modelos basados exclusivamente en variables racionales o contextuales. Este hallazgo es plenamente coincidente con los resultados del presente estudio, donde las variables emocionales (satisfacción, seguridad, esperanza, miedo) demostraron un poder predictivo muy superior a las variables demográficas (H1 confirmada). La coincidencia entre un estudio latinoamericano y el presente trabajo sugiere que la dimensión emocional es transversal en contextos de consumo de la región, independientemente del producto o servicio analizado.

Sobre el trade-off entre predicción y diagnóstico: Alantari (2022), en una tesis doctoral que analiza 260,489 reseñas de 25,241 productos mediante un estudio comparativo de métodos de *machine learning* para análisis de sentimientos, identificó un trade-off fundamental: los modelos basados en redes neuronales ofrecen las predicciones más precisas, mientras que los modelos de tópicos (como LDA) proporcionan mejores capacidades diagnósticas. Este hallazgo es relevante para interpretar los resultados del presente estudio: *Random Forest* demostró alta capacidad predictiva (AUC = 0.887), pero su naturaleza de *caja negra* limita la interpretación causal. Sin embargo, el análisis de importancia de variables (MeanDecreaseGini) permitió identificar la jerarquía de predictores (satisfacción y seguridad como los más influyentes), mitigando parcialmente esta limitación y ofreciendo un equilibrio entre predicción e interpretabilidad.

## 6.2. Implicaciones Teóricas (*Scientia*)

Reconfiguración de la jerarquía de variables: Los resultados demuestran que las variables emocionales poseen mayor poder predictivo que las demográficas en la compra de seminuevos. La satisfacción y seguridad emergieron como predictores clave, duplicando la importancia de variables como el deseo. Esto cuestiona los modelos lineales que tratan las emociones como constructos de igual peso y sugiere que, en contextos de alta incertidumbre, los consumidores priorizan emociones de *confirmación* (satisfacción) y *protección* (seguridad) sobre las de *excitación* (deseo, esperanza). Este hallazgo amplía el modelo de procesamiento dual de Kahneman (2011), proponiendo una subjerarquía dentro del Sistema 1 donde dominan las emociones vinculadas a evaluación post-compra y mitigación de riesgos, lo que es consistente con la noción de "*emociones utilitarias*" propuesta por Ladhari, Souiden y Dufour (2017).

La multicolinealidad emocional como fenómeno estructural: El estudio revela que la alta colinealidad entre variables emocionales (ej.  $r = 0.83$  entre satisfacción y seguridad) no es un obstáculo metodológico, sino una característica estructural que indica que las emociones operan como sistema sincronizado. La aplicación exitosa de , que maneja inherentemente esta multicolinealidad, valida este planteamiento. Esto implica que futuros modelos para bienes complejos deben abandonar el supuesto de independencia entre predictores emocionales y adoptar frameworks basados en redes o sistemas, como sugieren Tinoco-Egas, Juanatey-Bóga y Martínez-Fernández (2019) en su análisis de generación de emociones en la intención de compra.

Hacia una teoría integrada del consumo sostenible: La preeminencia de seguridad y confianza como facilitadores conecta directamente con los principios de economía circular y **ODS12**. Los resultados indican que la principal barrera para el consumo sostenible son las barreras emocionales ligadas a percepción de riesgo, no el precio o calidad. Esto sugiere la necesidad de una teoría integrada que combine economía circular con hallazgos de economía conductual y psicología de las emociones, donde la adopción de modelos sostenibles esté inversamente relacionada con barreras emocionales de riesgo y directamente relacionada con facilitadores de confianza y seguridad. Esta perspectiva es consistente con los hallazgos de Borg *et al.* (2024) sobre la necesidad de entender motivaciones profundas de los consumidores para priorizar comportamientos de consumo responsable.

### **6.3. Implicaciones Prácticas (Praxis)**

Los hallazgos proporcionan un *mapa de navegación emocional* para distribuidores y formuladores de políticas:

**Reducir miedo y construir seguridad:** Priorizar la disipación proactiva de miedos mediante certificaciones transparentes, periodos de prueba extendidos y garantías integrales. Esto es especialmente relevante en el contexto latinoamericano, donde Lovera y Cardinale (2023) identificaron que el análisis de sentimientos en plataformas digitales puede revelar percepciones de riesgo no evidentes en encuestas tradicionales.

**Maximizar satisfacción:** Rediseñar el proceso de compra con transparencia total en historial del vehículo, financiamiento simplificado y post-venta proactiva. La satisfacción no es solo un resultado, sino un predictor activo de lealtad futura.

Segmentar por perfiles emocionales: Dado el bajo poder predictivo de variables demográficas, segmentar según perfiles como *cautelosos* (requieren mensajes de seguridad) o *entusiastas* (responden a *esperanza y orgullo*). Esta estrategia es consistente con las recomendaciones de Alantari *et al.* (2022) sobre la necesidad de alinear métodos analíticos con objetivos de negocio específicos.

Desarrollar estándares y sellos de confianza: Crear un sello de calidad nacional estandarizado para vehículos seminuevos, abordando directamente la necesidad de seguridad y reduciendo la asimetría de información. Esta propuesta se alinea con las conclusiones de Franco-Santacruz y Serrano-Orellana (2025) sobre la importancia de factores contextuales en la predicción de compra en retail físico.

### **6.4. Limitaciones y Futuras Investigaciones**

Es importante reconocer las limitaciones del estudio: el uso de una muestra no probabilística en una sola región y el diseño transversal restringen la generalización estadística y el establecimiento de causalidad. Futuras investigaciones deberían: (1) emplear diseños longitudinales que permitan observar la evolución de las emociones a lo largo del proceso de compra; (2) replicar el estudio en diversas regiones de México y Latinoamérica para validar la consistencia de los hallazgos; (3)

explorar el uso de técnicas de *deep learning* (como **LSTM**) para comparar su rendimiento con *Random Forest* en este tipo de datos, siguiendo la línea de Lovera y Cardinale (2023); y (4) incorporar análisis de texto de reseñas online para complementar los datos de encuesta, como sugieren Alantari *et al.* (2022).

A pesar de estas limitaciones, este trabajo establece un precedente metodológico sólido y multidisciplinario para explorar los drivers emocionales que pueden acelerar la transición hacia sistemas de consumo más sostenibles en Latinoamérica.

## 7. CONCLUSIÓN

Esta investigación se planteó como pregunta central: ¿cómo influyen las emociones en la decisión de compra de vehículos seminuevos en México, y de qué manera un enfoque de *machine learning* puede contribuir a predecir dicho comportamiento en el marco de la innovación de proceso y el consumo sostenible? Los hallazgos presentados permiten responder a esta pregunta de manera concluyente.

Este estudio establece como nuevo conocimiento que las emociones —particularmente la satisfacción con el proceso y la seguridad percibida— son predictores robustos y jerárquicamente superiores a las variables demográficas en la compra de seminuevos en México. Mediante un modelo *Random Forest* validado con 1,000 casos en Aguascalientes (**AUC-ROC** = 0.887), se demostró que: (1) las variables emocionales explican la mayor parte de la varianza en la propensión a la recompra, superando significativamente a los modelos lineales tradicionales (**H1** y **H4** confirmadas); (2) las emociones operan como un sistema estructurado en dos clusters —afecto positivo y negativo— que actúan interdependientemente (**H2** confirmada); y (3) la satisfacción y la seguridad emergen como los predictores más influyentes, muy por encima del deseo o la esperanza (**H3** confirmada).

Estos hallazgos integran, por primera vez en el contexto latinoamericano, la economía conductual, la psicología de las emociones y la ciencia de datos para abordar un problema de consumo sostenible. Desde una perspectiva aplicada, se demuestra que las principales barreras para la adopción de la economía circular en el sector automotriz son de naturaleza emocional —específicamente el miedo y la desconfianza estructural— y no técnica o económica.

A partir de estos resultados, se delinearán tres líneas para la agenda de investigación futura:

**Línea 1.** Estudios longitudinales y causales. Futuras investigaciones deberían emplear diseños longitudinales que permitan observar la evolución de las emociones a lo largo del proceso de compra (pre, durante y post) y establecer relaciones causales más sólidas.

**Línea 2.** Replicación y ampliación contextual. Se requiere replicar el estudio en otras regiones de México y Latinoamérica para validar la consistencia de los hallazgos en distintos contextos socioculturales y económicos, incorporando análisis comparativos entre países.

**Línea 3.** Exploración de métodos alternativos de *machine learning*. Dado que estudios recientes (Lovera y Cardinale, 2023) sugieren que modelos de *deep learning* como **LSTM** pueden superar a *Random Forest* en ciertos contextos, se recomienda explorar su rendimiento en datos tabulares de encuestas, así como incorporar análisis de texto de reseñas online (Alantari, 2022).

En síntesis, esta investigación confirma el papel central de las emociones en decisiones de consumo de alto valor y aporta un marco metodológico innovador para el consumo sostenible. Contribuye directamente al **ODS12** y sienta bases para estrategias comerciales y políticas públicas que, al restaurar la confianza, aceleren la transición hacia una movilidad circular y sostenible en México y la región.

## 8. REFERENCIAS

- Aconauto. (2025, febrero 21). *Latin American car sales grew 7.8% in 2024, with Brazil and Mexico as market drivers*. *Aftermarket International*. Recuperado el 3 de marzo de 2026, de <https://www.aftermarketinternational.com/en/news/latest-news/9227-latin-american-car-sales-grew-7-8-in-2024-with-brazil-and-mexico-as-market-drivers.html>
- Aguilar-Cruz, P. D., & Campos-Sánchez, A. (2024). Fostering sustainable development through social innovation: The role of cultural values in entrepreneurial intentions. *Scientia et PRAXIS*, 4(08), 96–126. <https://doi.org/10.55965/setp.4.08.a4>
- Alantari, H. J. (2022). *An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews* [Tesis doctoral, University of California, Irvine]. eScholarship. <https://escholarship.org/uc/item/7q62b9b8>
- Ariely, D. (2016). *Payoff: The hidden logic that shapes our motivations*. Simon & Schuster. <https://www.simonandschuster.com/books/Payoff/Dan-Ariely/TED-Books/9781501120046>
- Asociación Mexicana de Distribuidores de Automotores. (2024). *Reporte de mercado interno automotor al último mes de 2024*. [https://www.amda.mx/wp-content/uploads/2412\\_Reporte\\_Mercado\\_Automotor.pdf](https://www.amda.mx/wp-content/uploads/2412_Reporte_Mercado_Automotor.pdf)
- Barenca-Sotelo, C. U., Maciel-Arellano, M. D. R., & Larios-Rosillo, V. M. (2026). Hacia una innovación de proceso de evaluación docente sostenible: Revisión sistemática de

- aplicaciones de IA, ciencia de datos y procesamiento de lenguaje natural. *Scientia et PRAXIS*, 6(11). <https://doi.org/10.55965/setp.6.11.a1>
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt. <https://www.amazon.com.mx/How-Emotions-Are-Made-Secret/dp/0544133315>
- Borg, K., Macklin, J., Kaufman, S., & Curtis, J. (2024). Consuming responsibly: Prioritising responsible consumption behaviours in Australia. *Cleaner and Responsible Consumption*, 12, 100181. <https://doi.org/10.1016/j.clrc.2024.100181>
- Fargues, M., Kadry, S., Lawal, I. A., Yassine, S., & Rauf, H. T. (2023). Automated analysis of open-ended students' feedback using sentiment, emotion, and cognition classifications. *Applied Sciences*, 13(4), 2061. <https://doi.org/10.3390/app13042061>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Franco-Santacruz, K. F., & Serrano-Orellana, B. J. (2025). Predicción de compra basada en estados emocionales y factores contextuales en retail físico usando *machine learning*. *Sociedad & Tecnología*, 8(S2), 327–341. <https://doi.org/10.51247/st.v8iS2.647>
- Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Instituto Nacional de Estadística y Geografía. (2023). *Inventario Nacional de Emisiones de Gases y Compuestos de Efecto Invernadero (INEGI)*. <https://www.inegi.org.mx/rnm/index.php/catalog/1013>
- Instituto Nacional de Estadística y Geografía. INEGI. (2024). *Registro administrativo de la industria automotriz de vehículos ligeros (RAIAVL)*. <https://www.inegi.org.mx/datosprimarios/iavl/>
- Jain, V. K., Dahiya, A., Tyagi, V., & Sharma, P. (2023). Development and validation of scale to measure responsible consumption. *Asia-Pacific Journal of Business Administration*, 15(5), 795–814. <https://doi.org/10.1108/APJBA-12-2020-0460>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux. <https://us.macmillan.com/books/9780374533557/thinkingfastandslow/>
- Ladhari, R., Souiden, N., & Dufour, B. (2017). The role of emotions in utilitarian service settings: The effects of emotional satisfaction on product perception and behavioral intentions. *Journal of Retailing and Consumer Services*, 34, 10–18. <https://doi.org/10.1016/j.jretconser.2016.09.005>
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, 66, 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Li, X., Wang, C., Li, D., Yang, D., Meng, F., & Huang, Y. (2024). Environmental regulations, green marketing, and consumers' green product purchasing intention. *Sustainability*, 16(20), 8987. <https://doi.org/10.3390/su16208987>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22. [https://cran.r-project.org/doc/Rnews/Rnews\\_2002-3.pdf](https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf)
- Loewenstein, G., O'Donoghue, T., & Bhatia, S. (2015). Modeling the interplay between affect and deliberation. *Decision*, 2(2), 55–81. <https://doi.org/10.1037/dec0000029>

- Lovera, F. A., & Cardinale, Y. (2023). Sentiment analysis in Twitter: A comparative study. *Revista Científica de Sistemas e Informática*, 3(1), e418. <https://doi.org/10.51252/rsos.v3i1.418>
- Murillo-López, F. J. (2025). Towards sustainable digital education: A predictive model for preventing social media addiction in university students. *Scientia et PRAXIS*, 5(10), 94–125. <https://doi.org/10.55965/setp.5.10.a4>
- Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Achem, J. A. G., Escamilla, J., & Hosseini, S. (2022). Towards teaching analytics: A contextual model for analysis of students' evaluation of teaching through text mining and *machine learning* classification. *Education and Information Technologies*, 27(3), 3891–3933. <https://doi.org/10.1007/s10639-021-10751-5>
- Organisation for Economic Co-operation and Development.OECD. (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD Publishing. <https://doi.org/10.1787/9789264304444-en>
- Organisation for Economic Co-operation and Development. OECD. (2024). *The future of the automotive value chain: Implications for FDI-SME linkages* (OECD SME and Entrepreneurship Papers No. 64). OECD Publishing. <https://doi.org/10.1787/cb730d65-en>
- Ortegón-Cortázar, L., Santucci, M., Iglesias-Pina, D., Acevedo-Duque, Á., & Méndez-Lazarte, C. (2025). Consumo sostenible y conciencia de las generaciones futuras en Latinoamérica. *RETOS. Revista de Ciencias de la Administración y Economía*, 15(30), 203–224. <https://doi.org/10.17163/ret.n30.2025.01>
- Peña-Torres, J. A. (2024). Towards an improved of teaching practice using sentiment analysis in student evaluation. *Ingeniería y Competitividad*, 26(2). <https://doi.org/10.25100/iyv.v26i2.13759>
- Pinzón-Castro, S. Y., & Maldonado-Guzmán, G. (2023). Open innovation effects in eco-innovation and business performance in Mexican manufacturing firms. *Scientia et PRAXIS*, 3(6), 1–19. <https://doi.org/10.55965/setp.3.06.a1>
- Posit Team. (2024). *RStudio: Integrated development environment for R* (Versión 2024.04.2) [Software]. Posit Software, PBC. <https://posit.co/>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77. <https://doi.org/10.1186/1471-2105-12-77>
- Santamaria-Velasco, C. A., Montiel-Méndez, O. J., & Montañez-Moya, G. S. (2024). Liderazgo transformacional y emprendimiento en estudiantes: Una vía hacia el desarrollo educativo sostenible. *Scientia et PRAXIS*, 4(08), 90–119. <https://doi.org/10.55965/setp.4.08.uady.a4>
- Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & O'Neill, S. (2022). Educational decision support system adopting sentiment analysis on student feedback. *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 377–383. <https://doi.org/10.1109/WI-IAT55865.2022.00062>
- Simon, H. A. (2016). *Administrative behavior: A study of decision-making processes in administrative organizations* (5th ed.). Free Press. (Original work published 1947).
- Thaler, R. H. (2018). *Misbehaving: The making of behavioral economics*. Penguin Books. <https://www.penguin.co.uk/books/179906/misbehaving-by-thaler-richard/9780241951224>

- Tinoco-Egas, R., Juanatey-Bóga, Ó., & Martínez-Fernández, V. A. (2019). Generación de emociones en la intención de compra. *Revista de Ciencias Sociales*, 25(3), 218–229. <https://www.redalyc.org/articulo.oa?id=28060161018>
- Villegas, J. C., Sánchez, J. L. M., & Rojas, E. (2023). Intención emprendedora en estudiantes universitarios. *Scientia et PRAXIS*, 3(6), 1–21. <https://scientiaetpraxis.amidi.mx/index.php/sp/article/view/65>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wei, T., & Simko, V. (2021). *corrplot: Visualization of a correlation matrix* (Version 0.92) [R package]. <https://CRAN.R-project.org/package=corrplot>
- Yu, S., Zhong, Z., Zhu, Y., & Sun, J. (2024). Green emotion: Incorporating emotional perception in green marketing to increase green purchase intentions. *Sustainability*, 16(12), 4935. <https://doi.org/10.3390/su16124935>
- Zhang, F., Chen, J., Tang, Q., & Tian, Y. (2024). Evaluation of emotion classification schemes in social media text: An annotation-based approach. *BMC Psychology*, 12, Article 503. <https://doi.org/10.1186/s40359-024-02008-w>
- Zhang, M. (2025). Decoding consumer behavior in the used car market: A machine learning approach to key decision factors. *ITM Web of Conferences*, 70, 04033. <https://doi.org/10.1051/itmconf/20257004033>

## Anexo 1. CONSENTIMIENTO INFORMADO

### TÍTULO DEL ESTUDIO:

“Emociones y toma de decisiones en la compra de automóviles seminuevos y usados en México”

INSTITUCIÓN: Universidad Autónoma de Aguascalientes

INVESTIGADOR RESPONSABLE: Dr Francisco Jacobo Murillo López

### CONSENTIMIENTO INFORMADO PARA PARTICIPANTES

Estimado(a) participante:

Se le invita a participar en un estudio académico cuyo propósito es analizar cómo las emociones influyen en la decisión de compra de automóviles seminuevos. Su participación es completamente voluntaria y confidencial.

Por favor lea cuidadosamente la siguiente información:

#### 1. Propósito del estudio

Comprender la relación entre emociones, percepción de riesgo y decisiones de compra en el mercado de autos seminuevos.

#### 2. Procedimiento

Si acepta participar, se le pedirá responder un cuestionario de aproximadamente 10 a 15 minutos o

#### 3. Riesgos y beneficios

No se prevén riesgos mayores a los de la vida cotidiana.

No recibirá compensación económica, pero su participación contribuirá a la investigación en marketing y comportamiento del consumidor.

4. Voluntariedad

Su participación es voluntaria. Puede retirarse en cualquier momento sin consecuencia alguna.

5. Confidencialidad

- No se recopilará información que lo identifique directamente (nombre, teléfono, correo).
- Sus respuestas serán codificadas con un número de participante.
- Los datos serán utilizados únicamente para fines académicos.

6. Uso de la información

Los resultados podrán ser publicados en artículos científicos, tesis o presentaciones académicas, sin revelar información personal o específica que permita su identificación.

7. Aceptación

Marque una opción:

Sí, acepto participar en este estudio.

No acepto participar.

Firma del participante (opcional): \_\_\_\_\_

Fecha: \_\_\_\_\_

Firma del investigador: \_\_\_\_\_

## **Anexo 2. CUESTIONARIO**

Se diseñó un cuestionario estructurado que incluye preguntas demográficas, económicas y emocionales. Las emociones fueron medidas mediante una escala tipo Likert del 1 al 5, basada en instrumentos validados como el Positive and Negative Affect Schedule (PANAS).

CUESTIONARIO: Emociones y Decisión de Compra de Automóviles Seminuevos

Sección 1. Datos generales

1. Edad:

\_\_\_\_ años

2. Género:

- Hombre
- Mujer
- Otro: \_\_\_\_\_

3. Salario mensual aproximado: \_\_\_\_\_

- Menos de \$20,000
- \$20,001 – \$40,000
- \$40,001 – \$60,000
- \$60,001 – \$80,000
- \$80,001 – \$100,000
- Más de \$100,000

Sección 2. Emociones durante el proceso de compra

Use la siguiente escala para responder:

1 = Totalmente en desacuerdo

2 = En desacuerdo

3 = Ni de acuerdo ni en desacuerdo

4 = De acuerdo

5 = Totalmente de acuerdo

Confianza

5. Confíe en la empresa o vendedor durante el proceso de compra.

6. Sentí que la información que recibí era confiable.

Entusiasmo

7. Me sentí emocionado(a) al buscar o evaluar el vehículo.

8. Sentí entusiasmo cuando encontré opciones que me interesaban.

Ansiedad

9. Me sentí ansioso(a) mientras evaluaba diferentes opciones.

10. La posibilidad de equivocarme me generó ansiedad.

Deseo

11. Sentía un fuerte deseo de obtener el auto que estaba viendo.

12. Quería avanzar rápidamente con la compra porque el auto me gustaba mucho.

Satisfacción

13. Me sentí satisfecho(a) con el proceso de compra.

14. Al finalizar la compra, sentí satisfacción con mi decisión.

Miedo

15. Tenía miedo de ser estafado(a) o engañado(a).

16. Sentí temor de comprar un vehículo con fallas ocultas.

Esperanza

17. Tenía expectativas positivas sobre el desempeño del vehículo.

18. Esperaba que mi compra mejorara mi vida o movilidad.

Orgullo

19. Sentí orgullo al considerar o comprar el automóvil.

20. La compra del vehículo me hizo sentir bien conmigo mismo(a).

Culpa

21. Me sentí culpable por gastar dinero en esta compra.

22. Experimenté culpa pensando en otros gastos o responsabilidades.

Confusión

23. Me sentí confundido(a) por la cantidad de información disponible.

24. Tuve dificultades para comparar opciones o entender las diferencias.

Impulsividad

25. Tomé decisiones rápidas durante la compra.

26. Avancé sin analizar todos los detalles que debía considerar.

Seguridad

27. Sentí seguridad en el vehículo que evalué o compré.

28. Me sentí seguro(a) de haber tomado una buena decisión.

### Anexo 3. CARTA PARA SOLICITAR ACCESO A EMPRESAS DE AUTOS SEMINUEVOS

[Aguascalientes, Ags. Méx.], [Fecha]

A quien corresponda:

Nombre de la empresa

Dirección

Ciudad

Estimados(as) señores(as):

Mi nombre es \_\_\_\_\_, profesor/investigador de \_\_\_\_\_.

Me permito solicitar su amable autorización para realizar una investigación titulada:

*“Emociones y toma de decisiones en la compra de automóviles seminuevos en México”.*

El estudio tiene como objetivo analizar cómo factores emocionales, cognitivos y de percepción de riesgo influyen en la adquisición de autos seminuevos. La información será obtenida de manera responsable mediante:

- Cuestionarios anónimos a clientes que hayan comprado un automóvil seminuevo en su establecimiento, o
- Entrevistas cortas a compradores recientes (opcional y si la empresa lo autoriza).

Compromisos del investigador:

1. No interferir en la operación o experiencia de venta.
2. No solicitar información personal de clientes.
3. No registrar datos sensibles o confidenciales.
4. Entregar resultados generales a la empresa si así lo solicita.
5. Garantizar anonimato total de participantes y de la empresa.

Agradezco de antemano su colaboración y quedo atento a resolver cualquier duda.

Atentamente,

Nombre

Teléfono

Correo electrónico

3. Protocolo de confidencialidad

#### **Anexo 4. PROTOCOLO DE CONFIDENCIALIDAD DE DATOS**

Estudio: *“Emociones y toma de decisiones en la compra de automóviles seminuevos”*

El presente documento establece las medidas que se adoptarán para garantizar la confidencialidad, seguridad y protección de los datos recabados en este estudio.

1. Naturaleza de los datos recolectados

Los datos consisten únicamente en:

- Variables sociodemográficas generales (edad, género, ingreso aproximado).
- Percepciones y emociones autoinformadas.
- Experiencia de compra.
- No se recogerán nombres, correos, teléfonos ni identificadores personales.

2. Codificación

Cada participante será identificado mediante un código alfanumérico (ej.: P001, P002).

Ningún archivo contendrá información personal asociada al código.

### 3. Almacenamiento seguro

- Los datos serán almacenados en archivos protegidos con contraseña.
- Solo el investigador responsable tendrá acceso.
- No se compartirán bases de datos sin anonimización.

### 4. Uso de datos

Los datos solo se utilizarán con fines académicos:

- artículos científicos
- reportes académicos
- presentaciones

Nunca se utilizarán para fines comerciales.

### 5. Eliminación de información

Una vez concluido el estudio, los datos serán resguardados por un periodo máximo de 2 años y posteriormente eliminados de forma definitiva.

Firma del investigador responsable: \_\_\_\_\_

Fecha: \_\_\_\_\_



This is an open access article distributed under the terms of the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)

# Scientia et PRAXIS

Vol. 06. No.11. Jan-Jun (2026): 59-85

<https://doi.org/10.55965/setp.6.11.a3>

eISSN: 2954-4041

## **Gender, Culture of Peace and Citizen Participation as Evidence-Based Social Innovation for Sustainable Urban Governance**

## **Género, cultura de paz y participación ciudadana como innovación social basada en evidencia para la gobernanza urbana sostenible**

**Tania Marcela Hernández-Rodríguez. ORCID [0000-0001-9946-0151](https://orcid.org/0000-0001-9946-0151)**

Universidad de Guadalajara, Centro Universitario de Guadalajara,

Guadalajara, Jalisco, México

e-mail: [tania.hernandez@cugdl.udg.mx](mailto:tania.hernandez@cugdl.udg.mx)

**César Omar Mora-Pérez. ORCID [0000-0002-3917-1144](https://orcid.org/0000-0002-3917-1144)**

Universidad de Guadalajara, Centro Universitario de Ciencias Económico-Administrativas,

Zapopan, Jalisco, México.

e-mail: [omar.mora@academicos.udg.mx](mailto:omar.mora@academicos.udg.mx)

**Keywords:** gender, culture of peace, citizen participation, neighborhood conflicts, sustainable community development

**Palabras Clave:** género, cultura de paz, participación ciudadana, conflictos vecinales, desarrollo comunitario sostenible

**Received:** Feb-3-2026; **Accepted:** May-12-2026

---

## ABSTRACT

**Context.** Gender inequalities in civic participation and conflict resolution persist in urban communities in the Guadalajara Metropolitan Area (ZMG), where a significant proportion of residents report insecurity or neighborhood conflicts, while participation in democratic processes remains limited. This situation reveals tensions among community coexistence, urban governance, and peacebuilding.

**Problem.** Although peace culture and citizen participation have gained relevance in studies of participatory governance and social innovation, the moderating role of gender in the relationship among peace culture, citizen participation, and perceptions of neighborhood conflicts remains underexplored. Therefore, it is pertinent to analyze how gender influences these relationships within urban communities.

**Purpose.** To analyze how gender moderates the relationship among peace culture, citizen participation, and neighborhood conflicts in urban communities of the ZMG, as a basis for designing social innovation interventions oriented toward sustainable development.

**Methodology.** A cross-sectional quantitative study was conducted in ten urban communities of the ZMG during March–April 2024, with residents ( $n = 229$ ). The analysis examined the relationship among peace culture, citizen participation, gender, and perceived neighborhood conflicts. Validated five-point Likert scales were applied. Data processing included exploratory factor analysis, Cronbach's alpha, Pearson correlations, and linear regression with interaction terms.

**Findings.** The scales showed high reliability ( $\alpha = .832-.931$ ). Peace, culture and citizen participation presented positive and significant correlations with the overall perception of conflicts. Gender significantly moderated the relationship between peace culture and conflict ( $B = .217, p = .043$ ), but not the relationship between citizen participation and conflict.

**Originality.** The study provides empirical evidence to redesign community practices related to participation, peace, and conflict management with a gender perspective, as a form of social and organizational process innovation in accordance with the Oslo Manual. It contributes to SDGs 5, 11, and 16.

**Conclusion.** Gender influences the relationship between peace culture and the perception of conflict; however, the cross-sectional design and geographic concentration limit the generalizability of the findings.

## RESUMEN

**Contexto.** Las desigualdades de género en la participación cívica y en la resolución de conflictos persisten en comunidades urbanas de la Zona Metropolitana de Guadalajara (ZMG), donde una proporción significativa de residentes reporta inseguridad o conflictos vecinales, mientras que la participación en procesos democráticos continúa siendo limitada. Esta situación evidencia tensiones entre convivencia comunitaria, gobernanza urbana y construcción de paz.

**Problema.** Aunque la cultura de paz y la participación ciudadana han adquirido relevancia en los estudios sobre gobernanza participativa e innovación social, el papel moderador del género en la relación entre cultura de paz, participación ciudadana y percepción de conflictos vecinales permanece insuficientemente explorado. Por ello, resulta pertinente analizar cómo el género influye en estas relaciones dentro de comunidades urbanas.

**Objetivo.** Analizar cómo el género modera la relación entre cultura de paz, participación ciudadana y conflictos vecinales en comunidades urbanas de la ZMG, como base para diseñar intervenciones de innovación social orientadas al desarrollo sostenible.

**Metodología.** Se desarrolló un estudio cuantitativo transversal en diez comunidades urbanas de la ZMG durante marzo-abril de 2024, con residentes urbanos ( $n = 229$ ). El análisis se centró en la relación entre cultura de paz, participación ciudadana, género y percepción de conflictos vecinales. Se aplicaron escalas validadas, con anclajes Likert de cinco puntos. El procesamiento incluyó análisis factorial exploratorio, alfa de Cronbach, correlaciones de Pearson y regresión lineal con términos de interacción para evaluar el efecto moderador del género.

**Hallazgos.** Las escalas mostraron alta confiabilidad ( $\alpha = .832-.931$ ). La cultura de paz y la participación ciudadana presentaron correlaciones positivas y significativas con la percepción general de los conflictos. El género moderó significativamente la relación entre cultura de paz y conflicto ( $B = .217, p = .043$ ), pero no la relación entre participación ciudadana y conflicto.

**Originalidad.** El estudio aporta evidencia empírica para rediseñar prácticas comunitarias de participación, paz y gestión de conflictos con enfoque de género, como innovación social y de proceso organizacional conforme al Manual de Oslo. Contribuye a los ODS 5, 11 y 16.

**Conclusión.** El género incide en la relación entre cultura de paz y percepción del conflicto; sin embargo, el diseño transversal y la concentración geográfica limitan la generalización.

## **1. INTRODUCCIÓN**

In recent decades, the analysis of contemporary social challenges has underscored the importance of examining power relations, participatory mechanisms, and conflict resolution from an integrated perspective. Gender, a culture of peace, citizen engagement, and community conflicts are not isolated phenomena; they are deeply interconnected elements that shape quality of life, substantive democracy, and social cohesion. Understanding how these variables interact in urban communities is essential for building stronger theoretical frameworks and designing more effective social interventions.

International scholarship has emphasized the need to integrate peacebuilding frameworks with participatory governance to address community-level conflicts. Galtung (1996) distinguishes between negative peace (understood as the absence of direct violence) and positive peace, which refers to the presence of social structures promoting justice, equality, and cooperation. Within this framework, conflicts are not necessarily dysfunctional; they can become opportunities for social transformation when managed through democratic dialogue and collective participation. This dynamic is well documented in urban contexts across Latin America, where neighborhood-level tensions frequently reflect deeper inequities in access to decision-making (Contreras-Ibáñez et al., 2019).

A gender perspective is indispensable for analyzing the structural inequalities that shape individual and collective experiences in social spaces. Women, historically excluded from public decision-making, face barriers that hinder their equitable participation and recognition as political actors. Incorporating gender into the analysis of social phenomena reveals these inequalities and supports the development of strategies for more just, inclusive societies. This is directly aligned with Sustainable Development Goal 5 (SDG 5: Gender Equality) and Goal 16 (SDG 16: Peace, Justice and Strong Institutions) of the United Nations 2030 Agenda.

The present study contributes to the literature on social innovation for sustainable development by examining how gender-sensitive governance dynamics shape peacebuilding, participation, and conflict perception in urban communities (Mejía-Trejo et al., 2025). Its originality lies in empirically testing a gender moderation model between culture of peace, citizen participation, and neighborhood conflict perception in urban communities of the GMA (a nexus that has not been directly examined in Latin American metropolitan settings) and in translating the results into

actionable social and process innovation, as defined by the Oslo Manual (OECD/Eurostat, 2018), for sustainable urban governance.

In terms of innovation type, this study aligns with organizational-process innovation as defined by the OECD/Eurostat (2018) it does not create a new product, but rather generates empirical evidence that supports redesigning participation practices, peacebuilding protocols, and conflict-management routines within urban community organizations. This framing situates gender-sensitive governance not as a normative ideal but as an actionable innovation input with measurable social impact, directly contributing to SDG 5, SDG 11, and SDG 16.

This study analyzes gender as a moderating variable in the relationship among the culture of peace, citizen participation, and neighborhood conflicts in urban communities, addressing a gap in empirical evidence in Latin American metropolitan settings. The specific objectives are: (a) to assess the level of association between a culture of peace, citizen participation, and neighborhood conflicts based on gender; and (b) to examine whether gender influences the strength or direction of these relationships. The central hypothesis states that gender significantly moderates the relationship between the culture of peace and perceptions of neighborhood conflict, with this link being stronger among men than among women. A secondary hypothesis proposes that gender moderates the relationship between citizen participation and conflict perception.

The article is organized as follows: Section 2 situates the study at global, international, national, and metropolitan levels. Section 3 reviews the theoretical literature across four axes and presents the measurement instrument and conceptual model. Section 4 describes the methodology. Sections 5 and 6 report results and discuss their theoretical and practical implications, respectively. Section 7 presents conclusions, limitations, and directions for future research.

## **2. CONTEXT**

This section situates the study across four levels of analysis (global, international, national, and metropolitan) to show how gender inequality, civic disengagement, and neighborhood conflict converge into a structural problem that calls for social innovation in urban governance.

## **2.1. Global**

At the global level, peacebuilding and participatory governance research consistently identify gender inequality as a structural barrier to sustainable community development. The United Nations (2015) 2030 Agenda establishes explicit targets linking peace (SDG 16), gender equity (SDG 5), and sustainable cities (SDG 11), recognizing that these challenges are inseparable in metropolitan environments. Cross-national evidence shows that countries with deeper gender gaps in civic participation also exhibit higher rates of community conflict and lower institutional trust (UNDP, 2022).

## **2.2. International and Latin American context**

At the international level, Latin American cities reproduce and intensify these dynamics: urban insecurity disproportionately affects women, while civic engagement mechanisms remain structurally biased toward male participation (Comisión Económica para América Latina y el Caribe [CEPAL] (2022)). Across the region, women's rates of political and community participation lag behind those of men by margins that persist even when controlling for education and socioeconomic status; CEPAL (2022) further documents that poverty and social exclusion (both more prevalent among women) reduce the capacity to engage in collective conflict resolution, reinforcing cycles of vulnerability that sustainable urban governance must directly address (UNDP, 2022).

## **2.3. National context**

Mexico, and particularly the state of Jalisco, presents a complex social landscape marked by insecurity, civic disengagement, and recurrent community conflicts. According to the Instituto Nacional de Estadística y Geografía (INEGI, 2024), between 62% and 75% of the population perceives their environment as unsafe, with a pronounced gender gap: women consistently report feeling more insecure than men. This pattern reflects broader structural inequalities that shape how different social groups experience and navigate urban life.

Civic participation rates in Jalisco reinforce this picture of social fragmentation. According to the Instituto Electoral y de Participación Ciudadana de Jalisco (IEPC, 2021), voter turnout in the 2021 local elections reached only 47.5% for state representatives and 47.84% for municipal

authorities. Participation in direct-democracy mechanisms was even lower: only 7% of the electorate voted in a state referendum on the fiscal pact, far below the 33% threshold required for a binding result.

## **2.4 Metropolitan context (GMA)**

Neighborhood conflicts are a pervasive reality in the Guadalajara Metropolitan Area (GMA). According to the Eighth Citizen Perception Survey on Quality of Life in Jalisco (Jalisco ¿Cómo Vamos? [JCV], 2022), between 72% and 91% of GMA residents have experienced community conflicts related to excessive noise, waste disposal, parking disputes, or pets. Of these conflicts, between 23% and 50% escalated to confrontations, including shouting, insults, or other verbal altercations. These data reflect fragile social cohesion and limited capacity for peaceful conflict resolution in everyday life.

Survey data from the Eighth Citizen Perception Survey on Quality of Life in Jalisco (JCV, 2022) reveals that 70% of the population reports little to no interest in politics or public affairs. Furthermore, only 36% are willing to sign a petition, 25% to share information on social media, and a mere 16% to participate in protests or demonstrations. These figures highlight a worrying trend of civic disengagement, directly undermining communities' capacity to resolve conflicts peacefully and inclusively.

Taken together, these indicators reveal a pattern of distrust, institutional disengagement, and limited civic participation that directly affects social cohesion and communities' capacity to manage conflicts peacefully and inclusively. This context makes it urgent to develop gender-aware strategies that foster a culture of peace, civic education, and equitable participation.

Read together, the data from the GMA trace a coherent causal pattern: pervasive insecurity disproportionately experienced by women, declining voter turnout, and widespread civic disengagement combine with frequent neighborhood conflicts to produce a fragile governance environment. This convergence (gender inequality, low participation and neighborhood conflict) defines the problem space that justifies the social and process innovation proposed in this study: empirical evidence to redesign gender-sensitive community governance for sustainable urban development.

### 3. LITERATURE REVIEW

This section reviews the state of the art across four theoretical axes: gender perspective, culture of peace, community conflicts, and citizen participation. The review draws on sources published predominantly between 2014 and 2025, retrieved from Scopus, Web of Science, Redalyc, and Google Scholar using the following search terms: gender AND peace culture, citizen participation AND urban conflict, gender moderation AND community governance, and social innovation AND sustainable development. Seminal works published before this window [Galtung (1996), Arnstein (1969), Pateman (1970)] are retained as theoretical anchors given their foundational role in structuring the conceptual model; each is complemented by recent empirical studies that demonstrate continued scholarly engagement with these frameworks. See **Table 1**

**Table 1. Literature Review Protocol: Search Strategy and Selection Criteria**

Element	Detail
Databases consulted	Scopus, Web of Science, Redalyc, Google Scholar
Search period	2014–2025 (seminal works retained regardless of year)
Search terms	<i>gender AND peace culture; citizen participation AND urban conflict; gender moderation AND community governance; social innovation AND sustainable development</i>
Initial results	340 records (approx. across all bases)
Inclusion criteria	Peer-reviewed articles and book chapters; thematic relevance to $\geq 1$ study axis; Spanish or English
Exclusion criteria	Grey literature without institutional affiliation; duplicates; works without identifiable methodology
Final documents reviewed	42
Thematic clusters identified	Gender perspective (n=12); Culture of peace (n=9); Community conflicts (n=8); Citizen participation (n=13)
Seminal works retained (>10 years)	Arnstein (1969), Pateman (1970), Barber (1984), Galtung (1996), Marshall (1997), Lederach (2003), Ansell & Gash (2008), Fraser (2009) — retained as foundational anchors; each complemented by post-2014 empirical studies

**Note:** Seminal works predating the search window are retained because their theoretical frameworks (positive peace, participatory democracy, feminist political theory, and collaborative governance) remain the canonical foundations for the constructs measured in this study. Their inclusion is standard practice in moderation-model research that requires theoretical grounding prior to operationalization. Source: Own elaboration.

#### *Gender Perspective*

Incorporating a gender perspective into the analysis of peace culture, citizen participation, and community conflicts is a crucial step toward understanding the power structures that have historically excluded women from the public sphere. In Mexico, the debate over women’s political

recognition has deep roots, as evidenced by the First Feminist Congress of Yucatán held in 1916, which highlighted fundamental debates about women's readiness to exercise political rights (Valles Ruiz, 2015; Peniche, 2017).

Over subsequent decades, feminist organizations such as the Mexican Feminist Council (CFM) and the United Front for Women's Rights (FUPDM) advocated for equal pay, maternity leave, access to employment, and voting rights. The constitutional reform of 1974 (Article 4) marked a milestone in establishing legal equality, reinforced a year later by the International Conference on Women in Mexico City (Rodríguez, 2015). Mexico's accession to CEDAW in 1980 formalized its international commitment to eliminating gender discrimination (United Nations, 1980).

The establishment of INMUJERES in 2001 institutionalized the gender perspective within government (INMUJERES, 2021). The General Law for Equality between Women and Men (2006) provided key definitional frameworks (Ministry of the Interior, 2013a). In 2018, Mexico achieved gender parity in Congress for the first time (INMUJERES, 2018), while the 2020 Protocol for Judging with a Gender Perspective issued by the Supreme Court (SCJN, 2020) further embedded gender analysis in institutional practice.

Contemporary feminist political theory further highlights the importance of analyzing participation and conflict through the lens of power and social inequalities. Fraser (2009) argues that democratic participation cannot be fully understood without considering structural asymmetries that influence individuals' capacity to impact public decision-making. Similarly, Young (2000) suggests that gender norms shape how people interpret conflict situations and engage in collective deliberation.

Taken together, this body of scholarship establishes that gender shapes civic life in structural and historically documented ways. A limitation common to these works, however, is the predominance of qualitative or descriptive approaches; quantitative evidence specifically testing the moderating effect of gender on the relationships among peace values, participation, and conflict perception in Latin American urban settings remains scarce. This gap justifies the moderation model proposed in the present study (H1 and H2). The growing attention to inclusive social innovation as a vehicle for reducing gender-based barriers in community contexts further reinforces the theoretical relevance of this inquiry (Carbajal-Silva y Aguilar-Cruz, 2025).

## ***Culture of Peace***

The culture of peace is an educational, social, and political framework that aims to transform human relations through dialogue, justice, and respect for human rights. Rooted in philosophical and religious traditions (from Confucian harmony and Christian theology to modern peace studies) it represents a fundamental shift toward nonviolent conflict resolution (Montserrat, 2010; Ruiz, 2014).

Galtung's (1996) theory of positive peace provides the foundational framework for understanding how social systems can prevent violence through structural justice and cooperative relationships. Lederach (2003) extends this by introducing conflict transformation, emphasizing long-term processes of dialogue, reconciliation, and social learning. These perspectives are particularly relevant in urban environments, where diverse interests and social identities interact in complex ways.

In Latin America, the pioneering work of Felipe MacGregor in Peru during the 1980s (culminating in the publication of "Culture of Peace" in 1986) established a systematic framework for educational and political transformation (Giesecke, 1999). Authors such as Jiménez-Bautista (2020), Andrade (2015), and Villalba (2016) have since emphasized that conflict should be viewed not as a problem but as an opportunity for positive transformation in social and educational relationships. Similarly, Martín, Extremera, and Martín (2018) demonstrate that school mediation can transform institutional culture by introducing democratic decision-making processes.

The literature on peace culture converges on the idea that violence is prevented through social structures rather than mere rule enforcement. What the empirical record has not addressed systematically is whether the relationship between peace-oriented values and conflict perception differs by gender — a question with direct implications for how peacebuilding interventions should be designed. Aguilar-Rosado and Campos-Sánchez (2024) offer a parallel lesson from the entrepreneurship domain: gender conditions not just access to resources but the very meaning actors assign to social action, suggesting that peace-culture studies would benefit from similar disaggregation.

### ***Community Conflicts***

Community conflicts are complex expressions of social tensions that arise when the interests of different groups clash. Far from being anomalies, they are natural components of social and territorial dynamics that should be managed with a focus on equity, participation, and dialogue (Reátegui & Álvarez, 2010). Research on social innovation in rural and indigenous communities has further illustrated how participatory and circular-economy strategies can transform conflict dynamics by fostering social cohesion and equity—an insight increasingly applicable to urban contexts (Mejía-Trejo y Mora-Pérez, 2025).

Pérez-Pereyra (2012) emphasizes that community relations structures must go beyond instrumental roles to ensure communities' rights to participation and self-determination. Trust, cultivated through consistent actions and active involvement of legitimate community representatives, is foundational to constructive conflict management (Ventura, 2010). Espinoza-Freire and Campuzano-Vásquez (2019) further highlight that citizen platforms must be built as ongoing processes involving both technical and community oversight.

Recent research indicates that local conflicts arise from competing interests, communication failures, and unequal access to decision-making; however, it still offers limited evidence on how these dynamics are perceived differently across genders in urban community settings. Rather than treating conflicts solely as governance failures, scholars suggest they can serve as signals of active civic engagement (Reátegui y Álvarez, 2010). This perspective is consistent with broader frameworks of democratic governance that view conflict as productive when managed through inclusive participatory mechanisms.

Despite these contributions, the field still lacks rigorous quantitative evidence on gender differences in conflict perception within urban residential settings. Most studies focus on ethnic or class-based conflict; gender as a moderating variable in everyday neighborhood disputes remains underexplored. This absence defines the empirical gap that the present study seeks to address through its interaction-term regression approach.

### ***Citizen Participation***

Citizen participation is a vital component of modern democracies; however, its effect on conflict perception is neither linear nor universal, and may vary according to gendered experiences of the public sphere. Marshall (1997) identifies three aspects of citizenship (civic, political, and social) noting that the political dimension involves rights to exercise political power directly or through elected representatives.

From a democratic theory perspective, Pateman (1970) argues that participatory mechanisms not only enhance decision quality but also cultivate more informed and engaged citizens. Barber (1984) similarly emphasizes that participatory democracy promotes a citizenry capable of tackling collective challenges. In Latin America, Montecinos (2006) documents the adoption of participatory budgeting as a mechanism that strengthens local democracy.

Arnstein's "Ladder of Participation" (1969) maps levels of participation from manipulation to citizen control, emphasizing that genuine, meaningful involvement requires structural support. Ansell and Gash (2008) further develop this idea through collaborative governance frameworks, showing how cooperation among public agencies, civil society, and private actors can address complex social issues. Martínez-Palacios (2017) underscores the necessity of integrating feminist and intersectional perspectives to ensure that women, minorities, and marginalized groups are heard and valued in decision-making.

The accumulated evidence supports a consistent finding: participation is structured by social position, and gender is one of the most consequential dimensions of that positioning. Yet whether gender moderates the statistical relationship between participation and conflict perception (rather than simply shaping participation rates) is a question that the literature has not directly examined. This research gap is the second hypothesis tested here. Mejía-Trejo and Aguilar-Navarro's (2022) work on inclusive community protocols in Mexico provides methodological precedent for operationalizing social innovation in contexts where participation and equity are tightly coupled.

#### **3.1. Measurement instrument design**

The measurement instrument was developed through a systematic review of validated scales for the three constructs: culture of peace, citizen participation, and neighborhood conflict perception. Items were selected to capture the theoretical dimensions identified in the literature review and

adapted to the specific socio-cultural context of the Guadalajara Metropolitan Area. The process of indicator design adhered to rigorous criteria for pertinence and social-impact relevance (Mora-Pérez, 2024), ensuring that each item reflected meaningful variation in the constructs of interest.

The final instrument comprised 28 items distributed across three latent constructs: Culture of Peace (7 items), Citizen Participation (9 items), and Neighborhood Conflicts (12 items). All items used a five-point Likert scale with verbal anchors ranging from 1 (strongly disagree) to 5 (strongly agree); conflict items were reverse-coded so that higher scores indicate greater perceived conflict, consistent with the negative theoretical direction of that variable.

Scales for culture of peace and citizen participation were adapted from instruments validated in previous Mexican and Latin American studies, with modifications to vocabulary and contextual referents reviewed by two subject-matter experts prior to application: one specialist in social measurement and survey design with doctoral-level training, and one researcher with fieldwork experience in urban communities of the GMA.

Both experts evaluated item clarity, contextual pertinence, and alignment with the construct definitions; their observations were incorporated through an iterative revision process until consensus on wording was reached. Structural validity was evaluated through exploratory factor analysis, and reliability was assessed via Cronbach's alpha, both conducted before proceeding with inferential analyses. See **Table 2**.

**Table 2. Measurement Instrument: Constructs, Items, and Operationalization**

Variable	Definition	Items	Scale	References
Culture of Peace	Level of agreement with practices of dialogue, respect, cooperation and non-violent conflict resolution	I agree that equality, freedom of expression, and the right to education should be promoted; I understand the importance of empathy and tolerance in building a culture of peace; I recognize conflicts or situations of violence in today's world; I know educational programs or initiatives that promote a culture of peace in my community or country; I actively participate in activities or projects that promote a culture of peace in my environment; I recognize how the media can influence public perception of peace culture and conflicts; I am informed about national and international laws, treaties, or policies related to promoting a culture of peace;	Likert 1–5	Galtung (1996); Lederach (2003); Jiménez-Bautista (2020); Martín et al. (2018)

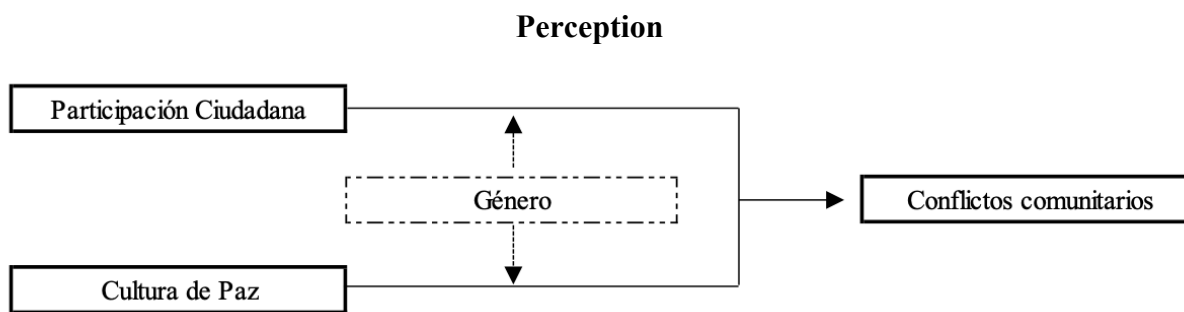
Variable	Definition	Items	Scale	References
Citizen Participation	Frequency or willingness to participate in community, deliberative or democratic actions	Promoting reforms to current legislation; Proposing effective public policies; Monitoring public spending and demanding accountability; Supporting investigative journalism; Organizing demonstrations and protests; Monitoring public management; Fostering a culture of civic reporting among citizens; Participating in public decision-making; Participating in citizen councils or commissions	Likert 1–5	Arnstein (1969); Pateman (1970); Barber (1984); Contreras-Ibáñez et al. (2019)
Neighborhood Conflicts	Perception of presence, intensity or frequency of community conflicts	Noise; Garbage thrown or burned by neighbors; Parking problems; Problems related to domestic animals; Gossip or misunderstandings; Disturbances caused by intoxicated individuals, drug users, or gangs; Conflicts in public or private transportation; Conflicts with neighbors' children; Parking disputes; Disputes with family members; Graffiti or vandalism on homes; Street vending conflicts	Likert 1–5	Andrade Correa (2015); Lederach (2003); Jalisco ¿Cómo Vamos [JCV] (2022)
Gender	Categorical variable used as a moderator	NA	Categorical	Fraser (2009); Young (2000); INMUJERES (2018, 2021)

Source: Own elaboration.

### 3.2. Conceptual model

The conceptual model proposes gender as a moderating variable that conditions the relationship between two independent variables—Culture of Peace and Citizen Participation—and the dependent variable, Neighborhood Conflict perception. This configuration reflects the theoretical argument that social identities, and gender in particular, shape how individuals perceive and respond to their community environment. Figure 1 illustrates the proposed relationships. Because the study employs a cross-sectional design, the model is understood to test statistical associations and moderation effects, not causal directionality; regression coefficients should therefore be interpreted as measures of association conditional on gender, not as causal estimates. See **Figure 1**.

**Figure 1. Conceptual Model: Moderating Effect of Gender on Community Conflict**



**Notes:** Solid arrows represent hypothesized direct relationships (IVs → DV); dashed arrows represent the proposed moderating paths (Gender × IV). This figure depicts the ex-ante theoretical model; empirical results are reported in Section 5.

Source: Own elaboration.

The model draws on three theoretical pillars: (a) Galtung’s positive peace framework, which situates peace culture as a structural quality of social environments; (b) democratic participation theory (Pateman, 1970; Barber, 1984), which links civic engagement to community resilience; and (c) feminist political theory (Fraser, 2009; Young, 2000), which emphasizes that structural gender inequalities mediate participation and conflict. The integration of these perspectives reflects the multidisciplinary approach aligned with SDG 5, SDG 11, and SDG 16.

#### 4. METHODOLOGY

This study employed a quantitative, cross-sectional survey design. A structured questionnaire was administered in person using a non-probabilistic convenience sampling strategy across ten communities in the GMA (Guadalajara, Zapopan, Tlaquepaque, Tonalá, and Tlajomulco) during March–April 2024. Communities were selected to ensure coverage of different socioeconomic strata and geographic zones within the metropolitan area; inclusion required that participants be adult residents (18 years or older) with at least six months of continuous residence in the community. Individuals who did not complete more than 20% of items or who failed an attention-check question were excluded. Of 251 questionnaires collected, 22 were discarded on these grounds, yielding 229 valid responses. The sample size exceeds the minimum recommended for exploratory factor analysis, commonly estimated at five to ten cases per item (Hair et al., 2019); in this study, 28 items × 5 = 140, while the final sample was n = 229 and provides sufficient statistical

power for multiple regression with interaction terms at the conventional  $\alpha = .05$  level (Cohen, 1988).

Gender was measured as a single categorical item with two response options (male / female), coded dichotomously (0 = female, 1 = male) for use in regression interaction terms. The sample comprised 121 women (52.8%) and 108 men (47.2%). The limitation of this binary operationalization — which does not capture non-binary or gender-diverse identities — is acknowledged and discussed in the conclusions.

Previously validated scales were culturally and contextually adapted to the GMA setting through item revision, adjustment of contextual wording, and internal consistency testing prior to inferential analysis. The operationalization of social impact variables, such as the culture of peace, citizen participation, and neighborhood conflict perception, adhered to rigorous indicator-design criteria (Mora-Pérez, 2024). **Table 3** summarizes the measurement structure.

**Table 3. Variables and Items**

Variable	Items	Sense	Type
Culture of Peace	7	+	Independent
Citizen Participation	9	+	Independent
Conflicts	12	-	Dependent

Source: Own elaboration.

Reliability analysis and exploratory factor analysis (EFA) were performed using IBM SPSS Statistics v.26. For the EFA, principal axis factoring was used as the extraction method, given the ordinal nature of the Likert data and the non-normality of several items confirmed via Kolmogorov-Smirnov tests ( $p < .05$ ). Oblimin oblique rotation was applied because theoretical correlation among the factors was expected. Factor retention followed Kaiser's criterion (eigenvalue  $> 1$ ) and was confirmed by a scree plot inspection. Missing data were minimal ( $< 1\%$  per item) and handled via listwise deletion. Interaction terms for the moderation models were constructed by mean-centering each continuous predictor before computing the product term with the dichotomous gender variable, in order to reduce multicollinearity (Cohen, 1988). All analyses were conducted after confirming data completeness and valid-range distributions through descriptive statistics and frequency distributions.

Two regression models were estimated for each independent variable. Model 1 assessed the direct relationship between the predictor (Culture of Peace or Citizen Participation) and the dependent variable (Neighborhood Conflicts). Model 2 introduced the Gender variable and the interaction term (Predictor  $\times$  Gender) to test for moderation.  $R^2$  change and ANOVA were used to evaluate model fit improvement.

**Ethical considerations.** Participation was voluntary and anonymous. Prior to data collection, all participants received a plain-language explanation of the study's purpose, their right to withdraw at any time without consequence, and the confidential treatment of their responses. Verbal informed consent was obtained from each participant, consistent with the non-sensitive, non-clinical nature of the study. No personally identifiable information was recorded. Data were stored in password-protected files accessible only to the research team and will be retained for five years in accordance with institutional guidelines. The study was conducted in adherence to the ethical principles of the Declaration of Helsinki (World Medical Association, 2013), for research with human participants and did not require formal ethics committee review, as it involved anonymous survey data with no foreseeable risk to respondents.

## **5. RESULTS**

This section presents the quantitative results from 229 valid surveys. It first verifies data completeness and instrument reliability, then examines factorial validity through KMO, Bartlett's test, and subscale reliability. Finally, it analyzes correlations and regression models to determine whether gender moderates the relationships between Citizen Participation, Culture of Peace, and Neighborhood Conflicts.

### **5.1 Descriptive statistics and reliability**

Initial descriptive analysis confirmed that data from all 229 surveys were complete and within valid ranges. The Cronbach's alpha for the full model was .871, indicating strong internal consistency. **Table 4** presents the case processing summary, and **Table 5** presents the overall reliability statistics.

**Table 4. Reliability Analysis – Case Processing Summary**

Cases	N	%
Valid	229	100.0
Excluded	0	.0
Total	229	100.0

Source: Own elaboration.

**Table 5. Overall Reliability Statistics**

Cronbach's alpha	Number of elements
.871	28

Source: Own elaboration.

## 5.2 Factorial analysis

Exploratory factor analysis yielded a Kaiser-Meyer-Olkin (KMO) value of .863 (excellent) and a statistically significant Bartlett’s test of sphericity ( $p < .001$ ), supporting the appropriateness of factor analysis for this dataset. Three main factors emerged, explaining 56.66% of the total variance, which is acceptable in social research. **Table 6** presents the KMO and Bartlett’s results.

**Table 6. KMO and Bartlett’s Test**

Kaiser-Meyer-Olkin sampling adequacy measure.	.863	
Bartlett's sphericity test	Approximate Chi-square	3989.933
	gl	378
	Sig.	0.000

Source: Own elaboration.

The rotated factor matrix grouped items clearly into the three theoretical constructs (Citizen Participation, Culture of Peace, and Neighborhood Conflicts), confirming structural validity. Individual reliability analyses for each subscale further corroborated this: Citizen Participation ( $\alpha = .911$ ), Culture of Peace ( $\alpha = .832$ ), and Neighborhood Conflicts ( $\alpha = .931$ ), all exceeding the recommended .70 threshold—see **Tables 7, 8, and 9** for these results.

**Table 7. Reliability Analysis for Citizen Participation**

Cronbach's alpha	Number of elements
.911	9

Source: Own elaboration.

**Table 8. Reliability Analysis for Culture of Peace**

Cronbach's alpha	Number of elements
.832	7

Source: Own elaboration.

**Table 9. Reliability Analysis for Neighborhood Conflicts**

Cronbach's alpha	Number of elements
.931	12

Source: Own elaboration.

### 5.3 Correlation analysis

A weak but significant positive correlation was found between Citizen Participation and Conflict ( $r = .159$ ,  $p = .016$ ). A moderate and significant positive correlation was found between Culture of Peace and Conflict ( $r = .202$ ,  $p = .002$ ). Citizen Participation and Culture of Peace were also significantly and positively correlated ( $r = .258$ ,  $p < .001$ ).

### 5.4. Correlations by gender

When variables were disaggregated by gender, a distinct pattern emerged. Among women ( $n = 121$ ), neither Citizen Participation nor Culture of Peace showed a statistically significant correlation with Conflict ( $r < .12$ ,  $p > .19$  for both); means were 3.41 (SD = 0.62) for Participation, 3.67 (SD = 0.58) for Culture of Peace, and 2.89 (SD = 0.71) for Conflict. Among men ( $n = 108$ ), both associations were significant: Citizen Participation correlated positively with Conflict ( $r = .26$ ,  $p = .007$ , 95% CI [.07, .43]) and Culture of Peace correlated positively with Conflict ( $r = .35$ ,  $p < .001$ , 95% CI [.17, .51]); means were 3.38 (SD = 0.65), 3.59 (SD = 0.61), and 2.94 (SD = 0.74), respectively. This gender differential constitutes preliminary evidence of a moderating effect, particularly in the culture of peace–conflict relationship.

### **5.5. Regression Analysis: Culture of Peace**

Model 1 showed that Culture of Peace significantly predicted conflict ( $B = .208$ ,  $p = .003$ ,  $R^2 = .048$ ). In Model 2, the introduction of the Gender interaction term rendered the main effect non-significant, while the interaction emerged as significant ( $B = .217$ ,  $p = .043$ ,  $R^2 = .065$ ). This increase in  $R^2$  from .048 to .065 confirms that gender moderates the relationship between the Culture of Peace and Neighborhood Conflicts.

### **5.6. Regression analysis: Citizen participation**

Model 1 showed that Citizen Participation significantly predicted Conflict ( $B = .151$ ,  $p = .020$ ,  $R^2 = .032$ ). In Model 2, the interaction term between Participation and Gender was not significant ( $B = .145$ ,  $p = .123$ ), and  $R^2$  increased only marginally (from .032 to .043). These results indicate that gender does not serve as a significant moderator in the participation–conflict relationship.

## **6. DISCUSSION**

The results reveal a complex and nuanced picture of how urban residents experience community conflict, shaped by their civic values, engagement in participation, and (critically) their gender. The overall finding that both Culture of Peace and Citizen Participation correlate positively with neighborhood conflict perception challenges conventional assumptions that these variables are straightforwardly conflict-reducing. A plausible interpretation is that higher levels of peace-oriented values and civic engagement heighten individuals' awareness of social tensions rather than suppressing them—a view consistent with democratic participation theories that highlight the constructive role of conflict in diverse societies (Barber, 1984; Pateman, 1970).

Citizen engagement can expose people to the tensions inherent in social pluralism, making conflicts more visible rather than creating them. Similarly, promoting a culture of peace may lead to greater recognition of injustice and community grievances, rather than their immediate resolution. These findings contribute to the literature on peace-building by underscoring that peace is not the absence of conflict but its conscious and participatory transformation. The moderating role of gender constitutes the most significant empirical contribution of this study. The finding that gender significantly moderates the culture of peace–conflict relationship (but not the participation–

conflict relationship) highlights the uneven ways in which peace values are experienced and expressed across genders.

Among men, higher scores on peace-oriented values were associated with greater conflict perception. One plausible account is that internalizing pacifist values heightens sensitivity to injustice (making tensions that might otherwise go unacknowledged more perceptible) though the cross-sectional design does not permit causal attribution and this interpretation warrants longitudinal testing.

One possible interpretation is that gendered socialization patterns influence how women perceive or report community conflict; however, this explanation should be corroborated through qualitative or longitudinal research. Women's lower conflict perception scores may reflect social roles oriented toward conciliation that led them to frame community tensions differently, or may reflect structural barriers that reduce their visibility in contexts where conflicts are measured. Either way, the absence of significant correlations in the female group calls for more targeted investigation rather than definitive conclusions about women's experiences of conflict.

The absence of gender moderation in the participation–conflict relationship suggests that civic engagement has a more direct and structurally shared effect: participation exposes both men and women to the tensions of social pluralism. This finding is consistent with Arnstein's (1969) conceptualization of participation as a structural position within the public sphere. However, it does not imply that participation is equally experienced across genders; rather, its relationship with conflict perception is statistically consistent.

### **6.1. Theoretical contributions (*Scientia*).**

From a theoretical perspective, this study makes several key contributions:

First, it provides empirical evidence that peace-oriented values are not gender-neutral constructs, challenging universalist models of peace culture and calling for gender-disaggregated conceptualizations of peacebuilding mechanisms.

Second, the study demonstrates that the relationship between civic engagement and community conflict perception is context-dependent and gender-conditioned. This contributes to theoretical frameworks of collaborative governance (Ansell y Gash, 2008) by showing that socially constructed identities mediate the effects of participation on conflict.

Third, the multidisciplinary integration of peace studies, feminist political theory, and participatory democracy theory advances an innovation-oriented understanding of community governance consistent with innovation management frameworks for sustainable development (Mejía-Trejo y Mora-Pérez, 2025). This positions gender equity not merely as a normative aspiration but as a functional input for sustainable community development aligned with SDG 5, SDG 11, and SDG 16.

## **6.2. Practical contributions (*Praxis*).**

The practical implications of this research are significant for the design of community policies aimed at fostering peace and civic participation. First, the finding that gender moderates the culture-of-peace–conflict relationship underscores the need for gender-differentiated approaches in community peacebuilding programs. Universal interventions that promote peace values without addressing gender dynamics risk generating unintended effects on perceptions of conflict.

Second, the study suggests that promoting citizen participation is a relatively gender-neutral strategy for enhancing conflict awareness, but gender-sensitive support structures must accompany it. The design of inclusive participation mechanisms—informed by circular-economy and social innovation frameworks for regional sustainable development (Mejía-Trejo y Mora-Pérez, 2025)—can help bridge existing structural gaps.

Third, local governments, NGOs, and academic institutions should collaborate in designing community programs that integrate peace education, civic training, and gender equity as mutually reinforcing pillars of sustainable urban development. The finding that only men show significant correlations among all three variables should not be interpreted as evidence that women are less relevant actors; rather, it highlights the need for research and policy tools that are sensitive to the different channels through which women engage with and manage community conflict.

## **7. CONCLUSION**

This section presents the conclusions of the study in three parts. First, the central research question is answered and both hypotheses are evaluated against the empirical evidence obtained. Second, the principal theoretical and practical findings are synthesized, highlighting the moderating role of gender and its differential expression across the variables studied. Third, the

limitations of the research are acknowledged and future lines of inquiry are proposed to advance the understanding of gender-moderated community dynamics in urban settings.

### **7.1. How to answer the question and research hypothesis.**

This study posed the central question of whether and how gender influences the relationships among a culture of peace, citizen participation, and neighborhood conflicts in urban communities. The evidence supports the first hypothesis: gender significantly moderates the relationship between culture of peace and conflict perception ( $B = .217, p = .043$ ), with the effect concentrated among men. The second hypothesis—that gender would also moderate the participation–conflict relationship—was not supported, as the interaction term failed to reach statistical significance ( $B = .145, p = .123$ ). These findings confirm that gender functions as a culturally embedded filter shaping how peace values—but not civic engagement per se—translate into perceptions of conflict.

### **7.2. Research findings**

Three principal findings emerge from this study. First, both the culture of peace and citizen participation are positively associated with perceptions of neighborhood conflict, a finding explained by the heightened awareness of tensions that accompanies civic engagement and peace-oriented values. Second, gender significantly moderates the culture of peace–conflict relationship, indicating that this construct operates differently across genders. Third, significant correlations between all three variables are observed only among men, pointing to important gendered differences in how community tensions are perceived and expressed. These findings provide empirical support for gender-sensitive, multidisciplinary approaches to sustainable community governance and reinforce the need to treat peacebuilding as a socially innovative process shaped by gendered experiences.

Specifically, the study's evidence contributes to SDG 5 by demonstrating that gender conditions how peace values translate into civic practice; to SDG 11 by identifying participation and peace culture as modifiable levers for urban conflict reduction; and to SDG 16 by grounding the call for inclusive institutions in quantifiable differences across gender groups. As a form of social and process innovation in the sense of the Oslo Manual (OECD/Eurostat, 2018), these results inform

the redesign of community governance interventions rather than merely documenting their current limitations.

### 7.3. Final research scope

This study has several limitations that should be acknowledged. The cross-sectional design precludes causal inference. Self-reported data may be affected by social desirability bias. The sample is geographically concentrated in the GMA, limiting transferability to other urban or rural contexts. Additionally, the binary operationalization of gender does not capture the full diversity of gender identities.

Future research should incorporate longitudinal designs to trace the evolution of gender-moderated dynamics over time. Comparative studies across regions and mixed-methods approaches that combine surveys with qualitative interviews would provide a richer understanding of the mechanisms linking gender, peace values, participation, and conflict. Expanding the operationalization of gender to include non-binary identities and integrating intersectional variables such as age, class, and ethnicity would further advance the explanatory power of this research agenda.

## 8. REFERENCES

- Aguilar-Rosado, T., & Campos-Sánchez, A. (2024). “Necesidad y oportunidad” como motivaciones para el emprendimiento femenino en Latinoamérica. *Scientia et PRAXIS*, 4(7), 31–57. <https://doi.org/10.55965/setp.4.07.a2>
- Andrade Correa, D. E. (2015). *Mediation and arbitration as alternative methods of conflict resolution in administrative contracts*. UCE. <https://www.dspace.uce.edu.ec/server/api/core/bitstreams/888f64b6-146c-4fb2-b05b-de897d8cffee/content>
- Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory*, 18(4), 543–571. <https://doi.org/10.1093/jopart/mum032>
- Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of Planners*, 35(4), 216–224. <https://doi.org/10.1080/01944366908977225>
- Barber, B. R. (1984). *Strong democracy: Participatory politics for a new age*. University of California Press. <https://www.ucpress.edu/book/9780520254480>
- Carbajal-Silva, A. A., & Aguilar-Cruz, P. D. (2025). Inclusive innovation in higher education: Evidence on entrepreneurship and sexual diversity. *Scientia et PRAXIS*, 5(10), 64–95. <https://doi.org/10.55965/setp.5.10.a3>

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum. <https://doi.org/10.4324/9780203771587>
- Comisión Económica para América Latina y el Caribe [CEPAL] (2022) *Panorama social de América Latina y el Caribe 2022*. Naciones Unidas. <https://repositorio.cepal.org/handle/11362/48518>
- Contreras-Ibáñez, C. C., Correa-Romero, F. E., & Patiño-Domínguez, H. A. (2019). Participación ciudadana y percepción del bienestar colectivo en comunidades urbanas. *Psicología y Salud*, 29(2), 197–208. <https://doi.org/10.25009/pys.v29i2.2597>
- Espinoza-Freire, E. E., & Campuzano-Vásquez, J. A. (2019). La participación ciudadana en la gestión local del Ecuador. *Revista de Ciencias Sociales*, 25(1), 90–101. <https://doi.org/10.31876/rcs.v25i1.27306>
- Fraser, N. (2009). *Scales of justice: Reimagining political space in a globalizing world*. Columbia University Press. <https://cup.columbia.edu/book/scales-of-justice/9780231146814>
- Galtung, J. (1996). *Peace by peaceful means: Peace and conflict, development and civilization*. SAGE. <https://doi.org/10.4135/9781446221631>
- Giesecke, M. (1999). Culture of peace and the teaching of history: Ecuador-Peru. En *Horizontes de negociación y conflicto* (pp. 303–315). FLACSO / DESCo. [http://flacso.edu.ec/docs/ecuaperu\\_giesecke.pdf](http://flacso.edu.ec/docs/ecuaperu_giesecke.pdf)
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage. [https://eli.johogo.com/Class/CCU/SEM/\\_Multivariate%20Data%20Analysis\\_Hair.pdf](https://eli.johogo.com/Class/CCU/SEM/_Multivariate%20Data%20Analysis_Hair.pdf)
- Instituto Electoral y de Participación Ciudadana de Jalisco. (2021). *Resultados del proceso electoral local 2021*. [https://www2.iepcjalisco.org.mx/tablero-electoral-2021/?page\\_id=1307](https://www2.iepcjalisco.org.mx/tablero-electoral-2021/?page_id=1307)
- Instituto Nacional de Estadística y Geografía. (2024, 24 de julio). *Encuesta Nacional de Seguridad Pública Urbana (ENSU)*. [https://www.inegi.org.mx/contenidos/programas/ensu/doc/ensu2024\\_junio\\_presentacion\\_ejecutiva.pdf](https://www.inegi.org.mx/contenidos/programas/ensu/doc/ensu2024_junio_presentacion_ejecutiva.pdf)
- INMUJERES. (2018). *First Congress of the Parity Union in Mexico*. <https://observatorio.inmujeres.gob.mx/mvc/view/public/index.html?q=MTA1>
- INMUJERES. (2021). *The gender digital divide: Another expression of inequality?* [http://cedoc.inmujeres.gob.mx/documentos\\_download/BA7N04%20VOBO\\_15072021.pdf](http://cedoc.inmujeres.gob.mx/documentos_download/BA7N04%20VOBO_15072021.pdf)
- Jalisco ¿Cómo Vamos? (2022). *Octava encuesta de percepción ciudadana sobre calidad de vida en Jalisco*. <https://jaliscocomovamos.org/infografias-epccv-2022/>
- Jiménez-Bautista, F. (2020). Conocer para comprender la violencia: Origen, causas y realidad. *Convergencia. Revista de Ciencias Sociales*, 58, 1–21. <https://doi.org/10.29101/rcs.v0i58.11544>
- Lederach, J. P. (2003). *The little book of conflict transformation*. Good Books. <https://www.goodbooks.com/products/the-little-book-of-conflict-transformation>
- Marshall, T. H. (1997). Citizenship and social class. En R. Goodin & P. Pettit (Eds.), *Contemporary political philosophy: An anthology* (pp. 291–319). Blackwell.

- <https://s3.us-west-1.wasabisys.com/p-library/books/b22baab2fe1f967c2346fb58c091c630.pdf>
- Martín, A. L., Extremera, P. G., & Martín, R. M. (2018). Educational mediation as a culture of peace. *Revista de Cultura de Paz*, 2, 125–145.  
<https://www.revistadeculturadepaz.com/index.php/culturapaz/article/view/26>
- Martínez-Palacios, J. (2017). Democratizing participation through feminism: The role of feminist subaltern counterpublics in expanding the Basque public sphere. *Spanish Journal of Political Science*, 43, 37–59. <https://doi.org/10.21308/recp.43.02>
- Mejía-Trejo, J., & Aguilar-Navarro, C. O. (2022). Valuando la evaluación: Protocolos comunitarios bioculturales, innovación de negocios circulares inclusivos y prospectivos en México. *Scientia et PRAXIS*, 2(3), 1–26. <https://doi.org/10.55965/setp.2.03.a1>
- Mejía-Trejo, J. & Mora-Pérez, C. O. (2024). Innovación de proceso: introducción del SROI como indicador de impacto social en los proyectos de los Centros Públicos de Investigación (CPI) de México. En C. O. Aguilar-Navarro (Coord.), *Diseño de indicadores de impacto social: Una discusión sobre su pertinencia en el desempeño de un centro público de investigación*. AMIDI.Biblioteca. <https://doi.org/10.55965/abib.9786076984512>
- Mejía-Trejo, J., Mora-Pérez, C. O. & Aguilar Navarro, C.O. (2025). La economía circular como estrategia de innovación para el desarrollo sostenible regional en México. En A. L. Moreno Ortiz, O. Aguilar Juárez, J. A. Medina Meléndez, & C. O. Aguilar Navarro (Coords.), *Innovación social y desarrollo sostenible en comunidades rurales e indígenas: Experiencias del sector agrario*. Ed. AMIDI / CIATEJ.  
<https://doi.org/10.55965/abib.9786072687547>
- Mejía-Trejo, J., y Mora-Pérez, C. O. (2025). *Administración de la innovación para el desarrollo sostenible: Fundamentos estratégicos*. AMIDI.Biblioteca.  
<https://doi.org/10.55965/abib.9786075818375>
- Ministry of the Interior. (2013a). *General Law for Equality between Women and Men*. Cámara de Diputados.  
[https://www.diputados.gob.mx/LeyesBiblio/ref/lgimh/LGIMH\\_ref03\\_14nov13.pdf](https://www.diputados.gob.mx/LeyesBiblio/ref/lgimh/LGIMH_ref03_14nov13.pdf)
- Montecinos, E. (2006). Decentralization and democracy in Chile. *Revista de Ciencia Política (Santiago)*, 26(2), 191–208. <https://www.scielo.cl/pdf/revcipol/v26n2/art11.pdf>
- Montserrat Antonin, M. (2010). *Intercultural mediation in Catalonia's health system* [Doctoral thesis, Universitat Rovira i Virgili].  
<https://www.tdx.cat/bitstream/handle/10803/8442/TESIS.pdf?sequence=1&isAllowed=y>
- OECD/Eurostat. (2018). *Manual de Oslo: Guía para la recogida, notificación y uso de datos sobre innovación* (4.ª ed.). OECD/Eurostat.  
[https://www.oecd.org/en/publications/oslo-manual-2018\\_9789264304604-en.html](https://www.oecd.org/en/publications/oslo-manual-2018_9789264304604-en.html)
- Pateman, C. (1970). *Participation and democratic theory*. Cambridge University Press.  
<https://www.cambridge.org/core/books/participation-and-democratic-theory/>
- Peniche, P. (2017). *Rita Cetina, La Siempreviva and the Literary Institute for Girls: A cradle of Mexican feminism, 1846–1908*. INERHM.  
<https://inehrm.gob.mx/work/models/inehrm/Resource/1484/1/images/RitaCetina.pdf>
- Pérez Pereyra, A. (2012). *Concepts and tools for the strategic management of community relations in extractive industries*. EOI.  
<https://www.eoi.es/blogs/artemioperez/files/2012/07/Herramientaspara-lagesti%C3%B3nestrat%C3%A9gica-de-relacionescomunitarias.pdf>

- Reátegui, J., & Álvarez, J. (2010). *Diálogo, dos años después: Perú: Estado y conflicto social*. ONDS-PCM. [http://onds.pcm.gob.pe/wp-content/uploads/2014/09/DIALOGO\\_ONDS.pdf](http://onds.pcm.gob.pe/wp-content/uploads/2014/09/DIALOGO_ONDS.pdf)
- Rodríguez, J. (2015). *Estado y transparencia: Un recorrido por la filosofía política*. INAI. [https://www.idaip.org.mx/archivos/formatos/Promocion\\_Vinculacion/Cuadernillo%2004%20B.pdf](https://www.idaip.org.mx/archivos/formatos/Promocion_Vinculacion/Cuadernillo%2004%20B.pdf)
- Ruiz Ortiz, C. (2014). Theology and culture of peace. *Cultura de Paz*, 20, 16–24. <https://doi.org/10.5377/cultura.v20i62.1466>
- SCJN. (2020). *Protocol for judging with a gender perspective*. Supreme Court of Justice of the Nation. <https://www.scjn.gob.mx/derechos-humanos/sites/default/files/protocolos/archivos/2020-11/Protocolo%20para%20juzgar%20con%20perspectiva%20de%20g%C3%A9nero%20%28191120%29.pdf>
- United Nations. (1980). *CEDAW: Convention on the Elimination of All Forms of Discrimination against Women*. SCJN. [https://www.scjn.gob.mx/igualdad-degenero/sites/default/files/cedaw/archivos/2021-11/convencion\\_discriminacion.pdf](https://www.scjn.gob.mx/igualdad-degenero/sites/default/files/cedaw/archivos/2021-11/convencion_discriminacion.pdf)
- United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development* (A/RES/70/1). <https://sdgs.un.org/2030agenda>
- United Nations Development Programme. (2022). *Human Development Report 2021/2022: Uncertain times, unsettled lives*. UNDP. <https://hdr.undp.org/content/human-development-report-2021-22>
- Valles Ruiz, R. (2015). First feminist congress of Mexico. En INERHM, *Historia de las mujeres en México*. INERHM. [https://www.uaeh.edu.mx/investigacion/productos/6834/primer\\_congreso\\_feminista\\_de\\_mexico- los\\_primeros\\_pasos.pdf](https://www.uaeh.edu.mx/investigacion/productos/6834/primer_congreso_feminista_de_mexico- los_primeros_pasos.pdf)
- Ventura, J. E. (2010). *La relación entre empresa y familia para la reducción de la pobreza: Negocios locales en un entorno rural* [Tesis doctoral, Universitat Internacional de Catalunya]. [https://www.tesisenred.net/bitstream/handle/10803/9174/VENTURA\\_Tesis\\_Doctoral\\_2010\\_11\\_08.pdf](https://www.tesisenred.net/bitstream/handle/10803/9174/VENTURA_Tesis_Doctoral_2010_11_08.pdf)
- Villalba, J. (2016). Curricular practices for positive coexistence and a culture of peace. *Itinerario Educativo*, 68, 131–146. <https://dialnet.unirioja.es/descarga/articulo/6290697.pdf>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>
- Young, I. M. (2000). *Inclusion and democracy*. Oxford University Press. <https://global.oup.com/academic/product/inclusion-and-democracy-9780198297550>



This is an open access article distributed under the terms of the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)

# Scientia et PRAXIS

Vol.06. No.11. Jan-Jun (2026): 86-118

<https://doi.org/10.55965/setp.6.11.a4>

eISSN: 2954-4041

## **Extended CAITIZEN: A PLS-SEM Study of Sustainable AI-Assisted Citizenship Innovation**

### **CAITIZEN extendido: estudio PLS-SEM de la innovación en ciudadanía sostenible asistida por IA**

**Juan Mejía-Trejo** ORCID [0000-0003-0558-1943](https://orcid.org/0000-0003-0558-1943)

Centro Universitario de Ciencias Económico Administrativas (CUCEA)

Universidad de Guadalajara (UdeG), Zapopan, Jalisco, México

e-mail: [jmejia@cucea.udg.mx](mailto:jmejia@cucea.udg.mx)

**Keywords:** ai-assisted sustainable citizenship, innovation for sustainable development, critical artificial intelligence literacy, pls-sem, higher education.

**Palabras Clave:** ciudadanía sostenible asistida por ia, innovación para el desarrollo sostenible, alfabetización crítica en inteligencia artificial, pls-sem, educación superior.

**Received:** Feb-13-2026; **Accepted:** Jun-24-2026

---

## ABSTRACT

**Context.** Artificial intelligence is transforming higher education by reshaping learning, creativity, decision-making, and civic participation. This study examines **CAITIZEN—Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked Formation**—as an extended model for validating AI-assisted sustainable citizenship as an innovation for sustainable development, aligned with **SDG4** and **SDG9**.

**Problem.** Although the original **CAITIZEN** model was qualitatively grounded as an ethical–cognitive–social framework, its explanatory and predictive capacity had not been empirically tested. AI education still prioritizes efficiency, automation, and technical adoption, with limited evidence on how critical AI literacy, ethics, data justice, human–AI collaboration, and metacognitive prompting contribute to sustainable AI-assisted citizenship.

**Purpose.** This study validates the extended **CAITIZEN** model through **PLS-SEM** by examining how **Critical Artificial Intelligence Literacy (CAIL)** enables **Ethical Awareness and Responsibility (EAR)**, **Awareness of Fairness and Data Justice (AFDJ)**, **Human–AI Creative Collaboration (HAIC)**, and **Metacognitive Transparency in Prompting Practices (MTPP)**, and how these capacities predict **CAITIZEN**.

**Methodology.** This research builds on a previous qualitative phase conducted in Guadalajara, Jalisco, Mexico, during July–December 2025, and complements it with an explanatory-predictive quantitative design using **SmartPLS 4.1.1.8** to assess reflective constructs and predictive relevance through **PLSpredict**.

**Theoretical and Practical Findings.** Results show that **CAIL** significantly predicts **EAR**, **AFDJ**, **HAIC**, and **MTPP**, confirming its role as foundational antecedent. **AFDJ**, **HAIC**, and **MTPP** significantly predict **CAITIZEN**, whereas **EAR** does not show a direct effect. Predictive relevance is confirmed because all **Q<sup>2</sup>\_predict** values for **CAITIZEN** indicators are positive and all **PLS-LM RMSE** differences favor **PLS-SEM**.

**Originality.** The study transforms the qualitative **CAITIZEN** model into an empirically validated explanatory-predictive structure.

**Conclusions and Limitations.** The extended **CAITIZEN** model provides a measurable framework for responsible AI education and sustainable innovation. Limitations include non-probabilistic sampling, cross-sectional design, and student sample.

## RESUMEN

**Contexto.** La inteligencia artificial transforma la educación superior al reconfigurar el aprendizaje, la creatividad, la toma de decisiones y la participación cívica. Este estudio examina **CAITIZEN —Ciudadanía Asistida por Inteligencia Artificial para una Formación Sostenible, Ética y en Red—** como modelo extendido para validar la ciudadanía sostenible asistida por **IA** como innovación para el desarrollo sostenible, alineada con el **ODS4** y el **ODS9**.

**Problema.** Aunque el modelo **CAITIZEN** original fue fundamentado cualitativamente como marco ético–cognitivo–social, su capacidad explicativa y predictiva no había sido probada empíricamente. La educación en **IA** aún prioriza eficiencia, automatización y adopción técnica, con evidencia limitada sobre cómo alfabetización crítica en **IA**, ética, justicia de datos, colaboración humano–**IA** y metacognición en prompts contribuyen a dicha ciudadanía.

**Propósito.** Este estudio valida el modelo **CAITIZEN** extendido mediante **PLS-SEM**, examinando cómo **CAIL** habilita **EAR**, **AFDJ**, **HAIC** y **MTPP**, y cómo estas capacidades predicen **CAITIZEN**.

**Metodología.** Este estudio parte de una investigación cualitativa previa realizada en Guadalajara, Jalisco, México, durante julio–diciembre de 2025, y lo complementa con un diseño cuantitativo explicativo-predictivo mediante **SmartPLS 4.1.1.8** para evaluar constructos reflectivos y relevancia predictiva mediante **PLSpredict**.

**Hallazgos teóricos y prácticos.** Los resultados muestran que **CAIL** predice significativamente **EAR**, **AFDJ**, **HAIC** y **MTPP**, confirmando su papel como antecedente fundacional. **AFDJ**, **HAIC** y **MTPP** predicen significativamente **CAITIZEN**, mientras **EAR** no muestra efecto directo. La relevancia predictiva se confirma porque todos los valores **Q<sup>2</sup>\_predict** son positivos y las diferencias **PLS-LM RMSE** favorecen a **PLS-SEM**.

**Originalidad.** El estudio transforma el modelo cualitativo **CAITIZEN** en una estructura explicativa-predictiva validada empíricamente.

**Conclusiones y limitaciones.** El modelo **CAITIZEN** extendido ofrece un marco medible para educación responsable en **IA** e innovación sostenible. Sus limitaciones incluyen muestreo no probabilístico, diseño transversal y muestra estudiantil.

## 1. INTRODUCTION

Artificial intelligence (AI) transforms higher education by reshaping learning, creativity, decision-making, civic participation, and institutional innovation. This article analyzes AI-assisted sustainable citizenship as a formative condition of university students who interact with intelligent systems and require critical, ethical, creative, and responsible competencies for sustainable development (Córdova-Esparza, 2025; Mejía-Trejo, 2025b).

In this study, **CAITIZEN** is retained as the name of the original model proposed by Mejía-Trejo (2025<sup>a</sup>), and its meaning is specified as Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked formation. This formulation preserves the focus on AI-assisted sustainable citizenship while making explicit its formative, ethical, and relational scope. The model includes five dimensions: Critical Artificial Intelligence Literacy (**CAIL**), Ethical Awareness and Responsibility (**EAR**), Awareness of Fairness and Data Justice (**AFDJ**), Human–AI Creative Collaboration (**HAIC**), and Metacognitive Transparency in Prompting Practices (**MTPP**), which frame citizenship in higher education (Mejía-Trejo, 2025<sup>a</sup>).

The problem is that AI education privileges efficiency, automation, and technical adoption, while limited evidence explains how these dimensions predict AI-assisted sustainable citizenship. This gap matters because universities require models connecting AI literacy, ethics, fairness, creativity, and metacognition with transformation, civic agency, learning, and responsibility (Southworth et al., 2023).

The solution is **CAITIZEN**, a structural model that analyzes how **CAIL**, **EAR**, **AFDJ**, **HAIC**, and **MTPP** predict AI-assisted sustainable citizenship. Its originality lies in transforming a qualitative ethical–cognitive–social framework into an empirically testable model. According to the Oslo Manual, it is posed as social and conceptual innovation for sustainable development, aligned with **SDG 4** and **SDG 9**, because it strengthens quality education and innovation-oriented transformation (OECD & Eurostat, 2005, 2018; United Nations, 2015).

The study integrates education, AI, ethics, data justice, creativity, metacognition, innovation, and sustainability. Research question: To what extent do **CAIL**, **EAR**, **AFDJ**, **HAIC**, and **MTPP** predict AI-assisted sustainable citizenship among higher education students? (Hair et al., 2022).

## 2. CONTEXT

This section contextualizes the environment in which AI-assisted sustainable citizenship develops and relates it to the empirical validation of **CAITIZEN**. The discussion is organized at three levels: global, international, and national-local. These levels show how artificial intelligence in higher education is no longer only a technological issue, but also an educational, ethical, civic, and innovation-for-sustainable-development challenge.

### 2.1. Global level

At the global level, artificial intelligence has reshaped education, work, innovation, and civic participation. UNESCO established that teachers and learners require AI competencies grounded in human-centered values, ethical principles, and responsible pedagogical use, which connects directly with SDG 4 on quality education (Miao & Cukurova, 2024). The World Economic Forum also observes that AI literacy is becoming a core future-work competence, because institutions and individuals face accelerated changes in skills, productivity, and innovation ecosystems (World Economic Forum, 2025). Likewise, the OECD reports a persistent AI skills gap, where technological adoption grows faster than training, governance, and critical understanding (OECD, 2025). In this environment, the subject of study—AI-assisted sustainable citizenship among higher education students—emerges as a formative condition required for responsible participation in AI-mediated societies.

### 2.2. International level

At the international level, the state of the art shows that AI literacy, ethical awareness, fairness, data justice, human–AI collaboration, and metacognitive prompting are increasingly treated as interdependent competencies. AI literacy is understood as the capacity to understand, evaluate, and use AI systems critically, beyond operational skills (Ng et al., 2021). Ethical awareness is observed as necessary for responsible decision-making and accountability when students use AI tools (Kong & Zhu, 2025). Fairness and data justice are established as core concerns because algorithmic systems can reproduce bias, discrimination, and unequal access to opportunities (Demirchyan, 2025). Human–AI creative collaboration is analyzed as an emerging innovation process in which AI supports ideation, design, writing, and knowledge production (Rafner et al., 2025).

Metacognitive prompting is also faced as a new learning challenge because students must formulate, monitor, and evaluate their interaction with generative AI systems (Haidar et al., 2025). Therefore, the object of study—**CAITIZEN**—is justified as an empirical model that connects these dimensions and evaluates their predictive relationship with AI-assisted sustainable citizenship.

### **2.3. México**

In México, digital adoption has expanded, but gaps remain in critical, ethical, and responsible **AI** use. INEGI reported that internet use is widespread among the population, yet digital access does not necessarily imply algorithmic understanding, data governance awareness, or responsible use of **AI** in educational contexts (INEGI, 2023). At the institutional level, UNESCO and CANIETI promoted a model for ethical and responsible artificial intelligence in Mexican organizations, showing that the country faces the need to move from technological adoption toward trustworthy **AI** ecosystems (UNESCO & CANIETI, 2025). Within higher education, this challenge becomes more specific because university students are citizens in formation who increasingly use AI for learning, creativity, decision-making, and academic production.

The previous **CAITIZEN** qualitative study established a contextual precedent by analyzing **511** undergraduate and graduate students and identifying five dimensions: **CAIL, EAR, AFDJ, HAIC, and MTPP**. That study positioned **CAITIZEN** as a conceptual and social innovation aligned with sustainable development (Mejía-Trejo, 2025<sup>a</sup>). However, the explanatory and predictive capacity of the model remains quantitatively untested. Therefore, **CAITIZEN** is proposed as a necessary empirical extension to validate whether those dimensions predict AI-assisted sustainable citizenship in higher education, contributing mainly to **SDG4** and complementarily to **SDG9** through educational innovation and institutional transformation. This contextualization supports the methodological transition from qualitative interpretation to structural validation, consistent with multidisciplinary innovation for sustainable development in universities.

## **3. LITERATURE REVIEW**

The extended **CAITIZEN** model is grounded in the need to explain AI-assisted sustainable citizenship as a multidimensional process in which critical literacy enables ethical, fairness-oriented, collaborative, and metacognitive capacities. The qualitative **CAITIZEN** study previously

conceptualized **AI**-assisted sustainable citizenship as an ethical–cognitive–social system composed of Critical Artificial Intelligence Literacy (**CAIL**), Ethical Awareness and Responsibility (**EAR**), Awareness of Fairness and Data Justice (**AFDJ**), Human–**AI** Creative Collaboration (**HAIC**), and Metacognitive Transparency in Prompting Practices (**MTPP**) (Mejía-Trejo, 2025<sup>a</sup>). However, the qualitative model was not designed to test causal relationships among these dimensions. Therefore, the present study re-specifies the original framework into an extended structural model for quantitative validation. In this re-specification, **CAIL** is positioned as a foundational antecedent, while **EAR**, **AFDJ**, **HAIC**, and **MTPP** are specified as proximal capacities that directly contribute to **CAITIZEN**. **CAIL** represents the cognitive and critical foundation of the model. **AI** literacy has been defined as the capacity to understand, use, evaluate, and critically engage with artificial intelligence systems. Long and Magerko (2020) established **AI** literacy as a competence that extends beyond technical operation and includes conceptual understanding of **AI** systems. Ng et al. (2021) conceptualized **AI** literacy as a multidimensional construct involving knowledge, application, evaluation, and ethical awareness. Southworth et al. (2023) further argued that **AI** literacy should be embedded across higher education curricula, while Wang and Wang (2025) showed that critical **AI** literacy supports reflective evaluation of **AI**-assisted writing. Therefore, **CAIL** is not treated only as a direct predictor of **AI**-assisted citizenship, but as a foundational condition that enables students to recognize ethical implications, fairness concerns, collaborative possibilities, and metacognitive demands in **AI**-mediated environments.

The first relationship proposed in the model links **CAIL** to **EAR**. Ethical awareness in **AI** use requires more than general moral sensitivity; it requires understanding how **AI** systems operate, how outputs are produced, and how algorithmic decisions may affect users, institutions, and society. Responsible **AI** literature emphasizes accountability, transparency, human oversight, and moral responsibility as central principles of **AI** governance. Gunasekara et al. (2025) synthesized responsible **AI** principles and practices, while Papagiannidis et al. (2025) framed responsible **AI** governance as a mechanism for aligning **AI** systems with ethical and organizational responsibility. Kong and Zhu (2025) validated an **AI** ethical awareness scale for students, showing that ethical awareness can be developed through **AI**-based problem-solving. Thus, students with higher **CAIL** are expected to show stronger **EAR**. Hence, we have:

**H1a. CAIL positively and significantly predicts EAR.**

The second relationship links **CAIL** to **AFDJ**. Awareness of fairness and data justice requires understanding that **AI** systems may reproduce bias, discrimination, unequal visibility, and unfair decision-making when data, algorithms, and institutional uses lack transparency or accountability. Decker et al. (2025) emphasized procedural fairness and public engagement in algorithmic decision-making. Demirchyan (2025) highlighted the regulatory challenges of algorithmic fairness, while González-Argote et al. (2025) identified algorithmic bias and data justice as ethical challenges in **AI** systems. In educational contexts, Pham et al. (2025) showed that fairness is a relevant concern in machine learning software. Since recognizing unfairness requires critical understanding of **AI** systems, **CAIL** is expected to strengthen **AFDJ**. Hence, we have:

**H1b. CAIL positively and significantly predicts AFDJ.**

The third relationship proposes that **CAIL** enables **HAIC**. Human–**AI** creative collaboration requires users to understand **AI** as a co-creative tool whose outputs must be guided, evaluated, and redirected by human agency. Georgieva and Georgiev (2025) examined the role of generative **AI** in design creativity, while Rafner et al. (2025) analyzed agency in human–**AI** collaboration for image generation and creative writing. Salma et al. (2025) identified paradoxes in co-creative systems, emphasizing the need to preserve human intentionality. Wang et al. (2025) showed that human–**AI** co-creation varies according to design experience. These studies suggest that critical **AI** literacy allows users to collaborate with **AI** more intentionally, creatively, and responsibly. Hence, we have: **H1c. CAIL positively and significantly predicts HAIC.**

The fourth relationship links **CAIL** to **MTPP**. Prompting practices are not merely technical instructions; they involve metacognitive awareness, goal clarification, evaluation of **AI** responses, and iterative refinement. Haidar et al. (2025) found that metacognitive prompts improve writing skills and metacognitive awareness in **AI**-supported learning. Tsakeni et al. (2025) mapped the role of **AI** tools in scaffolding metacognition and learning, while Waaler et al. (2025) demonstrated the importance of prompt engineering for improving compliance with intended instructional goals. Since students need critical understanding of **AI** systems to formulate, adjust, and evaluate prompts responsibly, **CAIL** is expected to strengthen **MTPP**. Therefore, we propose:

**H1d. CAIL positively and significantly predicts MTPP.**

The second stage of the model specifies **EAR**, **AFDJ**, **HAIC**, and **MTPP** as direct predictors of **CAITIZEN**. **EAR** is expected to contribute to **AI**-assisted sustainable citizenship because

ethical responsibility guides the use of **AI** toward accountability, academic integrity, social responsibility, and the common good. Responsible **AI** governance emphasizes that ethical principles are essential for trustworthy, legitimate, and socially aligned AI adoption (Gunasekara et al., 2025; Papagiannidis et al., 2025; UNESCO & CANIETI, 2025). Hence, we have:

**H2. EAR positively and significantly predicts CAITIZEN.**

**AFDJ** is expected to predict **CAITIZEN** because sustainable citizenship in AI-mediated societies requires awareness of bias, fairness, inclusion, and data justice. Algorithmic systems increasingly shape access, opportunity, visibility, and decision-making. Therefore, recognizing and questioning unfair AI outcomes becomes a civic competence linked to equity, trust, and responsible participation in digital environments (Decker et al., 2025; Demirchyan, 2025; González-Argote et al., 2025; Pham et al., 2025). Therefore, we have:

**H3. AFDJ positively and significantly predicts CAITIZEN.**

**HAIC** is expected to predict **CAITIZEN** because **AI**-assisted citizenship is not limited to ethical caution; it also involves creative agency, collaborative problem-solving, innovation-oriented participation, and the capacity to use **AI** as a partner for knowledge production. Human–**AI** collaboration can enhance creativity when users maintain agency, intentionality, critical control, and reflective judgment over the process (Georgieva & Georgiev, 2025; Rafner et al., 2025; Salma et al., 2025; Wang et al., 2025). Therefore, we have

**H4. HAIC positively and significantly predicts CAITIZEN.**

**MTPP** is expected to predict **CAITIZEN** because responsible **AI** use requires self-regulation, transparency, and conscious interaction with generative systems. Students who reflect on how they formulate prompts, evaluate outputs, adjust instructions, and recognize the limits of **AI** are better positioned to use **AI** autonomously, ethically, and sustainably. Metacognitive prompting therefore supports reflective **AI** engagement and responsible digital participation (Haidar et al., 2025; Tsakeni et al., 2025; Waaler et al., 2025). Hence, we propose:

**H5. MTPP positively and significantly predicts CAITIZEN.**

Finally, the model assumes that **CAIL** has indirect effects on **CAITIZEN** through the proximal capacities of **EAR**, **AFDJ**, **HAIC**, and **MTPP**. This assumption is consistent with **AI** literacy research, which frames literacy not merely as knowledge, but as a capacity that enables ethical judgment, critical evaluation, creative collaboration, and reflective regulation. In this sense, **CAIL**

functions as the foundational literacy layer of the model, while the other dimensions represent applied mechanisms through which **AI**-assisted sustainable citizenship is enacted.

Together, these arguments support the extended **CAITIZEN** model as a quantitative re-specification of the qualitative framework proposed by Mejía-Trejo (2025<sup>a</sup>). The original model identified the key dimensions of AI-assisted sustainable citizenship, whereas the present model specifies their internal structural logic by positioning **CAIL** as the foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**. These proximal capacities, in turn, are expected to shape **AI-assisted sustainable citizenship**. Thus, the model advances from qualitative conceptualization to quantitative explanation, while framing **CAITIZEN** as a social and conceptual innovation in higher education, aligned mainly with **SDG 4** and complementarily with **SDG 9** (OECD & Eurostat, 2018; United Nations, 2015).

### **3.1. Measurement Instrument Design**

The measurement instrument was designed as a quantitative continuation and structural refinement of the qualitative **CAITIZEN** framework. The questionnaire operationalized six reflective constructs: **CAIL**, **EAR**, **AFDJ**, **HAIC**, **MTPP**, and **AI-assisted sustainable citizenship (CAITIZEN)**. The items were derived from the conceptual definitions of the original qualitative model and from the reviewed literature on AI literacy, responsible AI, algorithmic fairness, human–AI collaboration, metacognitive prompting, and sustainable citizenship.

A Likert-type scale was selected to measure students' agreement with statements related to critical AI literacy, ethical responsibility, fairness and data justice, creative collaboration, metacognitive prompting, and AI-assisted sustainable citizenship. The instrument was structured to support **PLS-SEM** analysis by representing each construct through observable reflective indicators. This design enabled the evaluation of reliability, convergent validity, discriminant validity, collinearity, path coefficients, effect sizes, explanatory power, and complementary indirect effects.

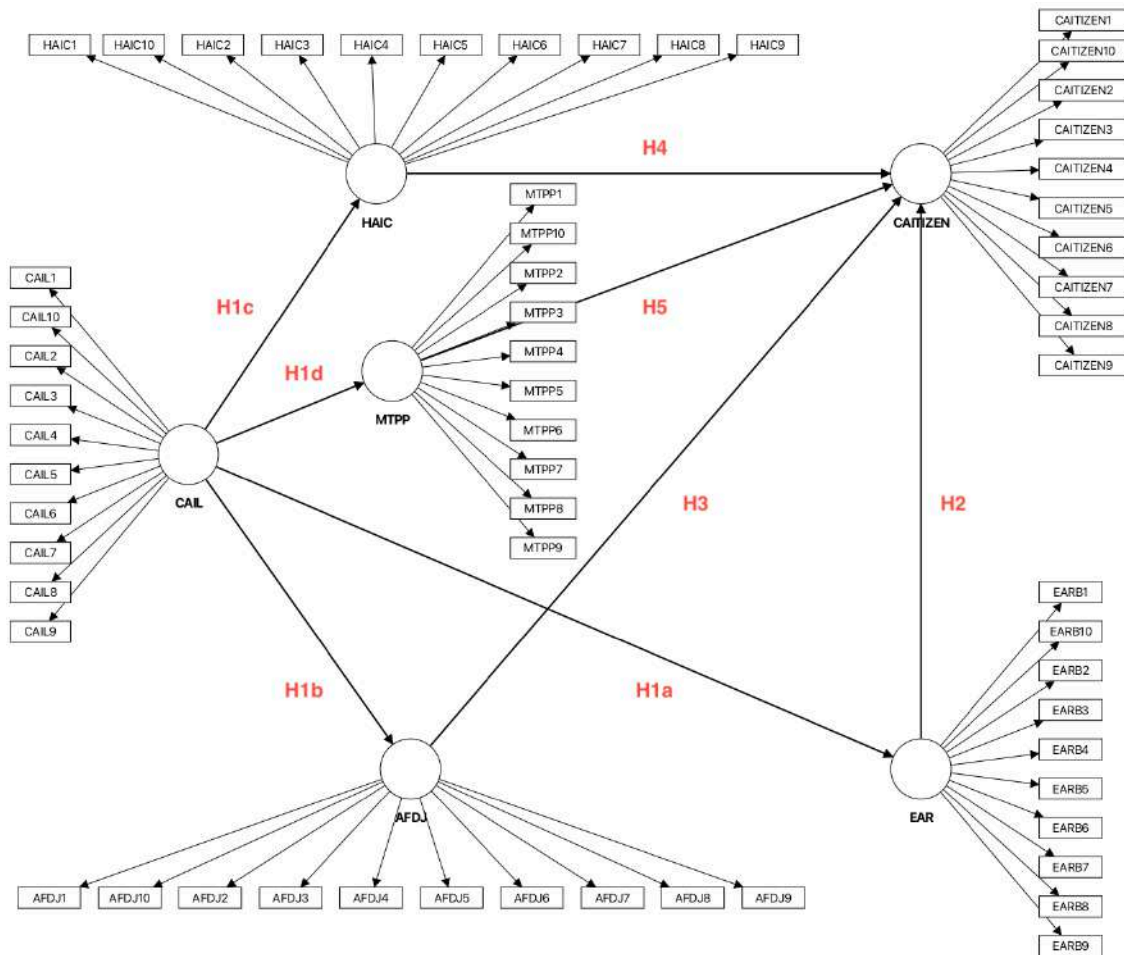
In line with the extended structural model, the instrument allowed **CAIL** to be examined as a foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**, while these four dimensions were specified as proximal predictors of **CAITIZEN**. Thus, the questionnaire translated the qualitative

dimensions of the **CAITIZEN** framework into measurable indicators, allowing the study to move from conceptual interpretation to empirical validation and structural explanation.

### 3.2. Conceptual Model

The **CAITIZEN** conceptual model is proposed *ex ante* as an extended structural model. It positions **CAIL** as the foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**, which in turn directly predict **AI-assisted sustainable citizenship (CAITIZEN)** among higher education students. Thus, the model does not replace the original qualitative **CAITIZEN** framework; it refines its explanatory structure and tests its predictive capacity through **PLS-SEM**. See **Figure 1**

**Figure 1. The extended CAITIZEN ex ante model**



**Note:** In this study, **CAITIZEN** is conceptually specified as *Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked formation*.

Source: Author's own elaboration based on the qualitative **CAITIZEN** model proposed by Mejía-Trejo (2025<sup>a</sup>).

This structure is justified because **CAIL** is specified as the foundational literacy condition, while **EAR**, **AFDJ**, **HAIC**, and **MTPP** represent proximal ethical, fairness-oriented, creative, and metacognitive capacities. **PLS-SEM** is appropriate because the study aims to test an explanatory and predictive model through reflective constructs, path coefficients, explained variance, effect sizes, and predictive relevance. Thus, the extended **CAITIZEN** model enables empirical validation of **AI-assisted sustainable citizenship** as a social and conceptual innovation in higher education.

#### **4. METHODOLOGY**

This study adopted a quantitative, explanatory-predictive research design aimed at empirically validating the extended **CAITIZEN** model. The model specifies Critical Artificial Intelligence Literacy (**CAIL**) as a foundational antecedent of Ethical Awareness and Responsibility (**EAR**), Awareness of Fairness and Data Justice (**AFDJ**), Human–AI Creative Collaboration (**HAIC**), and Metacognitive Transparency in Prompting Practices (**MTPP**), which in turn predict **AI-assisted Sustainable Citizenship (CAITIZEN)**. Given the growing role of artificial intelligence in higher education, the study examines how students’ critical AI literacy enables ethical, fairness-oriented, creative, and metacognitive capacities related to sustainable citizenship. As a quantitative continuation of the qualitative **CAITIZEN** framework, the methodological procedure translated its dimensions into measurable reflective constructs suitable for **PLS-SEM** analysis.

**Stage 1. Model Proposition.** The extended **CAITIZEN** model was proposed *ex ante* as a structural model in which **Critical Artificial Intelligence Literacy (CAIL)** functions as a foundational antecedent of four proximal capacities: **Ethical Awareness and Responsibility (EAR)**, **Awareness of Fairness and Data Justice (AFDJ)**, **Human–AI Creative Collaboration (HAIC)**, and **Metacognitive Transparency in Prompting Practices (MTPP)**. These four constructs, in turn, were specified as direct predictors of **AI-assisted Sustainable Citizenship (CAITIZEN)**. All constructs were operationalized as reflective constructs. The model assumes that AI-assisted sustainable citizenship among higher education students is shaped by ethical responsibility, fairness awareness, creative human–AI collaboration, and metacognitive prompting practices, while these capacities are enabled by critical AI literacy. Accordingly, eight direct hypotheses were proposed: four testing the effect of **CAIL** on **EAR**, **AFDJ**, **HAIC**, and **MTPP**, and four testing the effect of these proximal capacities on **CAITIZEN**.

**Stage 2. Data Analysis Technique. Partial Least Squares Structural Equation Modeling (PLS-SEM)** was selected because the objective was explanatory and predictive. PLS-SEM allowed the simultaneous assessment of the measurement model and the structural model. The measurement model was evaluated through indicator loadings, **Cronbach's alpha**, **composite reliability**, **rho\_A**, **average variance extracted (AVE)**, the **Fornell–Larcker criterion**, **HTMT**, and collinearity diagnostics. The structural model was assessed through **path coefficients**, **bootstrapping**, **coefficient of determination (R<sup>2</sup>)**, **effect sizes (f<sup>2</sup>)**, predictive relevance, and **PLSpredict**. This procedure made it possible to examine reliability, validity, explanatory power, effect magnitude, and predictive relevance. Data were analyzed using spreadsheet software for descriptive statistics and **SmartPLS** for **PLS-SEM** estimation.

**Stage 3. Study Subject and Sampling.** The subject of study consisted of undergraduate and graduate students who had interacted with artificial intelligence tools in academic, learning, creative, or decision-making activities. A non-probabilistic purposive sampling procedure was used because participants had to meet two inclusion criteria: being enrolled in higher education and having previous experience using AI tools. The study was not designed as a population census; it analyzed the complete and valid responses obtained from participants who met the inclusion criteria. Sample size adequacy was assessed according to **PLS-SEM** requirements, considering the five direct predictors of **CAITIZEN**, expected effect size, significance level, and desired statistical power.

**Stage 4. Data Collection.** Data were collected through an online questionnaire administered via Google Forms between July and December 2025. The questionnaire was distributed among undergraduate and graduate students of the University of Guadalajara, Jalisco, Mexico, through academic networks, institutional mailing lists, classroom groups, and online learning communities. This local-institutional scope is consistent with the study's purposive sampling strategy and supports the empirical validation of the **CAITIZEN** model within a higher education context. Participation was voluntary, anonymous, and without compensation. Before answering, participants were informed about the academic purpose of the study, the voluntary nature of participation, and the confidential treatment of responses. From approximately 800 invited participants, **600** accessed the questionnaire, and **511** complete and valid responses were retained after screening.

**Stage 5. Demographic Data.** Demographic data were analyzed to characterize the profile of the higher education students participating in the study. The variables considered included age, gender, marital status, educational level, type of higher education institution, and previous experience using artificial intelligence tools. These variables were not included as structural predictors in the **CAITIZEN** model; rather, they were used to contextualize the subject of study and describe the composition of the sample. Frequency and percentage distributions were calculated and organized in **Table 1**.

**Table 1. Demographic profile of the research sample**

Measure	Items	Frequency	Percentage	
Age	18-29	492	96.3%	100%
	30-39	10	2.0%	
	40-49	4	0.8%	
	50-59	5	1.0%	
Gender	Male	235	45.9%	100%
	Female	276	54.1%	
Marital Status	Single	500	97.8%	100%
	Married	11	2.2%	
Education Level	Undergraduate	499	97.6%	100%
	Master Degree	8	1.6%	
	Doctor Degree	4	0.8%	
Type Educational Institute or University	Public	461	90.2%	100%
	Private	50	9.8%	

Source: Own elaboration.

**Stage 6. Questionnaire Design and Data Preparation.** The questionnaire included six reflective constructs: **CAIL**, **EAR**, **AFDJ**, **HAIC**, **MTPP**, and **CAITIZEN**. Each construct was measured through ten Likert-type statements, resulting in **60** observable indicators. The items were formulated in first person and measured agreement with behaviors, perceptions, and competencies related to AI literacy, ethical responsibility, algorithmic fairness, creative collaboration, metacognitive prompting, and sustainable citizenship. A seven-point Likert scale was used, from 1 = Strongly disagree to 7 = Strongly agree. After data collection, the database was exported, coded, and cleaned. Missing values, duplicated entries, incomplete questionnaires, and inconsistent response patterns were reviewed. Only complete and valid responses were retained for descriptive and inferential analysis. The quantitative database corresponded to the closed-ended Likert-type section of the broader **CAITIZEN** data collection process. While the previous qualitative phase

analyzed open-ended responses, the present study focused on the closed-ended indicators suitable for **PLS-SEM** analysis.

## 5. RESULTS

This section presents the findings obtained from **Partial Least Squares Structural Equation Modeling (PLS-SEM)** using **SmartPLS 4.1.1.8**. The results are organized into three components: measurement model evaluation, structural model assessment, and empirical validation of the extended **CAITIZEN** model. The analysis first examines the reliability and validity of the reflective constructs. It then evaluates the structural relationships in which **CAIL** acts as a foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**, while these proximal capacities predict **AI-assisted Sustainable Citizenship (CAITIZEN)**. Finally, the empirical results are used to assess the explanatory and predictive capacity of the proposed model.

### 5.1. Measurement Model Evaluation

The reflective measurement model was assessed using established **PLS-SEM** criteria. The evaluation included internal consistency reliability, indicator reliability, convergent validity, discriminant validity, and collinearity diagnostics before proceeding to the structural model assessment.

#### 5.1.1. Reliability, Indicator Reliability, and Convergent Validity

Internal consistency reliability was evaluated through **Cronbach's alpha**, **rho\_A**, and **composite reliability**. The constructs **CAIL**, **EAR**, **AFDJ**, **HAIC**, **MTPP**, and **CAITIZEN** met the recommended reliability threshold of **0.70**, supporting adequate internal consistency.

Convergent validity was assessed through the **average variance extracted (AVE)**. The retained constructs achieved acceptable **AVE** values, indicating that each construct explained a sufficient proportion of the variance of its indicators. Indicator reliability was evaluated using **outer loadings**, and the retained indicators showed acceptable loadings and statistical significance after **Bootstrapping** with **[5,000]** subsamples.

Item purification was applied when indicators showed weak outer loadings, cross-loading concerns, or a negative effect on reliability, **AVE**, or discriminant validity. Indicators were

removed only when their exclusion improved the empirical quality of the construct without compromising its conceptual meaning. In total, **one indicator, CAIL2, was removed**, and the final measurement model retained **59 valid indicators**.

These results confirmed that the **CAITIZEN** measurement model achieved acceptable reliability and convergent validity. See **Table 2**.

**Table 2. Assessment of internal consistency, outer loadings, and convergent validity for the CAITIZEN measurement model**

Factor 1	Code	CAIL. Critical Artificial Intelligence Literacy. Students' capacity to understand, evaluate, verify, and responsibly use AI systems in academic, professional, and civic contexts Cronbach's Alpha ( $\geq 0.70$ ) = 0.884; rho_A (Dijkstra-Henseler's $\geq 0.70$ ) = 0.889; rho_c (Composite Reliability $\geq 0.70$ ) = 0.907; AVE ( $\geq 0.50$ ) = 0.521.	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references
Critical Artificial Intelligence Literacy	CAI L1	I understand how artificial intelligence systems transform data into patterns or predictions.	0.660 (19.115)	0.000	Long & Magerko (2020)
Critical Artificial Intelligence Literacy	CAI L2	I distinguish among the main types of artificial intelligence, such as predictive, generative, and autonomous AI, and their applications.	Indicator removed to improve the construct's AVE		Ng et al. (2021)
Critical Artificial Intelligence Literacy	CAI L3	I understand how the clarity and precision of a prompt influence AI-generated responses.	0.586 (15.182)	0.000	Waalder et al. (2025)
Critical Artificial Intelligence Literacy	CAI L4	I identify the technical limitations and margins of error of the AI models I use.	0.757 (31.206)	0.000	Southworth et al. (2023)
Critical Artificial Intelligence Literacy	CAI L5	I detect when an AI-generated output contains bias or unsupported inferences.	0.754 (29.560)	0.000	Wang & Wang (2025)
Critical Artificial Intelligence Literacy	CAI L6	I verify AI-generated information by comparing it with reliable academic or human sources.	0.715 (23.848)	0.000	Córdova-Esparza (2025)
Critical Artificial Intelligence Literacy	CAI L7	I recognize when a decision should be made by a person rather than by an algorithm.	0.744 (25.217)	0.000	Gunasekara et al. (2025)
Critical Artificial Intelligence Literacy	CAI L8	I analyze the ethical implications of using AI in teaching, research, and institutional management.	0.804 (42.441)	0.000	Miao & Cukurova (2024)

<b>Critical Artificial Intelligence Literacy</b>	CAI L9	I identify when an AI application may violate privacy, equity, or digital rights.	0.744 (29.890)	0.000	González-Argote et al. (2025)
<b>Critical Artificial Intelligence Literacy</b>	CAI L10	I promote the responsible, transparent, and traceable use of AI in my professional or academic environment.	0.710 (24.931)	0.000	Gunasekara et al. (2025)
Factor 2	Code	EAR. Ethical Awareness and Responsibility. Students' capacity to recognize ethical risks, assume responsibility, preserve academic integrity, protect privacy, and promote transparent and socially responsible AI use. Cronbach's Alpha ( $\geq 0.70$ ) = 0.911; rho_A (Dijkstra–Henseler's $\geq 0.70$ ) = 0.915; rho_c (Composite Reliability $\geq 0.70$ ) = 0.926 AVE ( $\geq 0.50$ ) = 0.555.	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references
<b>Ethical Awareness and Responsibility</b>	EAR 1	I explicitly declare when and how I use artificial intelligence in my academic work to maintain transparency and personal responsibility.	0.642 (19.683)	0.000	Kong & Zhu (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 2	Before using AI, I analyze possible ethical risks, such as bias, privacy, authorship, or misinformation, and how to prevent them.	0.718 (29.455)	0.000	Gunasekara et al. (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 3	I respect privacy and data protection when using AI tools by avoiding the disclosure of sensitive information.	0.730 (26.453)	0.000	González-Argote et al. (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 4	I avoid applying AI in ways that may cause harm, exclusion, or misinformation in my academic environment.	0.765 (31.081)	0.000	Córdova-Esparza (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 5	I encourage my peers to use AI ethically, transparently, and responsibly as a practice of academic integrity.	0.758 (33.419)	0.000	Kong & Zhu (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 6	I take responsibility for AI-generated outputs in my academic activities and correct errors derived from their use.	0.772 (32.162)	0.000	Papagiannidis et al. (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 7	I evaluate the social effects of the AI I use, considering equity, inclusion, and sustainability.	0.768 (33.737)	0.000	Córdova-Esparza (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 8	I recognize that academic products must preserve human authorship, using AI only as technical or creative support.	0.751 (29.612)	0.000	Wang & Wang (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 9	I reject using AI to plagiarize, falsify, or impersonate my own or others' academic work.	0.748 (26.849)	0.000	Kong & Zhu (2025)
<b>Ethical Awareness and Responsibility</b>	EAR 10	I promote transparency and the common good in the use of AI, ensuring that its benefits are shared equitably.	0.789 (40.308)	0.000	Papagiannidis et al. (2025)
Factor 3	Code	AFDJ. Awareness of Fairness and Data Justice. Students' capacity to identify algorithmic bias, recognize data-related inequalities, value transparency, and support fairness, accountability, and justice in AI systems. Cronbach's Alpha ( $\geq 0.70$ ) = 0.919; rho_A (Dijkstra–Henseler's $\geq 0.70$ ) = 0.921;	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references

		rho_c(Composite Reliability $\geq 0.70$ ) = 0.933; AVE ( $\geq 0.50$ ) = 0.581.			
<b>Awareness of Fairness and Data Justice</b>	AFDJ 1	I recognize that limited or homogeneous datasets may generate unfair algorithmic decisions.	0.684 (23.114)	0.000	González-Argote et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 2	I consider external auditing of algorithms necessary to detect and correct bias.	0.752 (28.681)	0.000	Decker et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 3	I value AI systems making public the criteria they use to make decisions.	0.773 (34.180)	0.000	Demirchyan (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 4	I believe that including cultural and gender diversity in development teams improves algorithmic fairness.	0.718 (26.522)	0.000	Pham et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 5	I identify the importance of continuously supervising AI models to prevent discrimination.	0.745 (31.297)	0.000	Demirchyan (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 6	I consider it essential for educational institutions to establish policies to evaluate algorithmic justice.	0.781 (31.711)	0.000	Decker et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 7	I observe that imbalance in data representation can affect a system's accuracy and impartiality.	0.812 (42.996)	0.000	Pham et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 8	I defend people's right to know the variables that influence an algorithm's decisions.	0.767 (27.094)	0.000	Decker et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 9	I consider transparency in the origin and processing of data to be key for equity in AI.	0.801 (39.930)	0.000	González-Argote et al. (2025)
<b>Awareness of Fairness and Data Justice</b>	AFDJ 10	I support the publication of institutional reports on detected biases and improvements implemented in educational algorithms.	0.779 (35.175)	0.000	Pham et al. (2025)
Factor 4	Code	HAIC. Human–AI Creative Collaboration. Students' capacity to use AI as a creative partner for ideation, problem-solving, experimentation, reflective improvement, and innovation-oriented knowledge production. Cronbach's Alpha ( $\geq 0.70$ ) = 0.942; rho_A (Dijkstra–Henseler's $\geq 0.70$ ) = 0.943; rho_c (Composite Reliability $\geq 0.70$ ) = 0.950; AVE ( $\geq 0.50$ ) = 0.656.	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references
<b>Human–AI Creative Collaboration</b>	HAIC 1	I use AI as a starting point to generate ideas that I later develop personally.	0.739 (27.678)	0.000	Georgieva & Georgiev (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 2	I combine my own contributions with AI suggestions to create more innovative outcomes.	0.812 (40.894)	0.000	Rafner et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 3	I analyze how AI responses influence my creative process.	0.795 (34.688)	0.000	Wang et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 4	I learn new ways of solving problems when interacting with AI.	0.829 (43.662)	0.000	Wang et al. (2025)

<b>Human–AI Creative Collaboration</b>	HAIC 5	I consider AI a creative partner that complements my human abilities.	0.801 (36.297)	0.000	Salma et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 6	Collaboration with AI inspires new perspectives or approaches that I would not have had alone.	0.784 (32.013)	0.000	Rafner et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 7	I use AI to experiment with different creative styles or solutions.	0.840 (47.686)	0.000	Georgieva & Georgiev (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 8	I reflect on how my work improves through interaction with AI.	0.817 (40.091)	0.000	Wang et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 9	I adjust AI proposals to align them with my own creative vision or purpose.	0.836 (50.771)	0.000	Salma et al. (2025)
<b>Human–AI Creative Collaboration</b>	HAIC 10	In creative projects, I combine my human decisions with AI’s generative capabilities to achieve original outcomes.	0.841 (50.968)	0.000	Rafner et al. (2025)
Factor 5	Code	MTPP. Metacognitive Transparency in Prompting Practices. Students’ capacity to plan, document, monitor, adjust, and reflect on prompting practices to improve learning, transparency, and responsible AI interaction. Cronbach’s Alpha ( $\geq 0.70$ ) = 0.918; rho_A (Dijkstra–Henseler’s $\geq 0.70$ ) = 0.923; rho_c (Composite Reliability $\geq 0.70$ ) = 0.932; AVE ( $\geq 0.50$ ) = 0.577.	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P1	I systematically record the prompts I use in order to learn from my own process.	0.680 (22.064)	0.000	Haidar et al. (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P2	I adjust my prompts when I detect that AI responses do not meet my learning objectives.	0.766 (30.472)	0.000	Haidar et al. (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P3	I analyze changes in AI responses to understand how my requests affect the results.	0.772 (31.853)	0.000	Tsakeni et al. (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P4	I reflect on the clarity and precision of my prompts before submitting them.	0.794 (39.200)	0.000	Haidar et al. (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P5	I document the reasoning I follow when modifying a prompt or interpreting a response.	0.714 (22.127)	0.000	Tsakeni et al. (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P6	I compare different versions of the same prompt to identify which one produces more useful results.	0.702 (20.172)	0.000	Waalder et al. (2025)
<b>Metacognitive Transparency</b>	MTP P7	I evaluate whether my prompting strategies contribute to a deeper understanding of the topic.	0.816 (44.317)	0.000	Tsakeni et al. (2025)

<b>in Prompting Practices</b>					
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P8	I plan the structure and purpose of the prompt before beginning my interaction with AI.	0.823 (47.701)	0.000	Miao & Cukurova (2024)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P9	I review the coherence between my initial objectives and the results obtained from AI.	0.776 (35.589)	0.000	Wang & Wang (2025)
<b>Metacognitive Transparency in Prompting Practices</b>	MTP P10	At the end of the activity, I reflect on what I learned from the process of interacting with AI.	0.741 (28.494)	0.000	Haidar et al. (2025)
Factor 6	Code	<b>CAITIZEN</b> . AI-assisted Sustainable Citizenship. Students' capacity to participate responsibly in AI-mediated societies by integrating critical literacy, ethics, fairness, creativity, metacognition, innovation, and social sustainability. Cronbach's Alpha ( $\geq 0.70$ ) = 0.910; rho <sub>A</sub> (Dijkstra-Henseler's $\geq 0.70$ ) = 0.912; rho <sub>c</sub> (Composite Reliability $\geq 0.70$ ) = 0.925; AVE ( $\geq 0.50$ ) = 0.553.	Outer loading ( <i>t-value</i> )	<i>p</i> - Value	Associated references
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 1	I identify myself as part of a society that coexists daily with artificial intelligence.	0.712 (25.165)	0.000	Mejía-Trejo (2025 <sup>a</sup> )
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 2	I consider it important to understand how AI influences my rights and responsibilities as a citizen.	0.740 (25.552)	0.000	Mejía-Trejo (2025 <sup>a</sup> )
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 3	I believe that everyone should participate in public decisions related to the use of AI.	0.719 (24.341)	0.000	Decker et al. (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 4	I am interested in AI being used transparently by governments and institutions.	0.742 (27.335)	0.000	Papagiannidis et al. (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 5	I think that access to AI knowledge is a form of citizen empowerment.	0.731 (29.056)	0.000	Southworth et al. (2023)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 6	I am concerned that AI may reinforce inequalities if inclusive policies are not implemented.	0.658 (18.382)	0.000	González-Argote et al. (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 7	I trust that informed citizenship can guide the ethical development of AI.	0.779 (34.155)	0.000	Gunasekara et al. (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 8	I am willing to learn continuously in order to adapt to an AI-driven society.	0.750 (24.031)	0.000	Córdova-Esparza (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 9	I believe that being a citizen in the AI era implies actively collaborating in its regulation and oversight.	0.795 (37.653)	0.000	Papagiannidis et al. (2025)
<b>AI-assisted Sustainable Citizenship</b>	CAITIZEN 10	I value the importance of combining technological innovation with collective responsibility and social sustainability.	0.795 (39.014)	0.000	Mejía-Trejo (2025 <sup>a</sup> )

Notes:

- **CAIL**. Critical Artificial Intelligence Literacy; **EAR**. Ethical Awareness and Responsibility; **AFDJ**. Awareness of Fairness and Data Justice; **HAIC**. Human–AI Creative Collaboration; **MTPP**. Metacognitive Transparency in Prompting Practices; **CAITIZEN**. AI-assisted Sustainable Citizenship. The items were conceptually adapted from the qualitative **CAITIZEN** model and supported by specialized literature. The associated references indicate conceptual support and do not imply that the items were directly copied from previous scales.
- Indicators with outer loadings between **0.40** and **0.70** were retained when reliability, **AVE**, and theoretical relevance supported their inclusion.
- **rho\_A**. Dijkstra–Henseler’s rho\_A is a reliability coefficient used to assess internal consistency in PLS-SEM. Values above 0.70 indicate acceptable reliability.
- **rho\_C**. Composite Reliability assesses the internal consistency of construct indicators. Values between 0.70 and 0.90 are considered satisfactory. Values above 0.95 may suggest indicator redundancy or common method bias (Hair et al., 2022).
- **AVE** (Average Variance Extracted). An **AVE** > 0.50 indicates that a construct explains more than 50% of the variance in its observed indicators, thereby supporting convergent validity (Fornell & Larcker, 1981).
- Indicators are according to Likert Scale 1-7 (1. Strongly Disagree; 2. Disagree; 3. Somewhat Disagree; 4. I do not know ; 5. Somewhat Agree; 6. Agree; 7. Strongly Agree).
- Regarding saturated model fit, the **SRMR** value was **0.048** below the recommended threshold of **0.08**, indicating acceptable approximate fit and supporting the continuation of the **PLS-SEM** assessment.

Source: Own elaboration using SmartPLS 4.1.1.8

### 5.1.2. Discriminant Validity

Discriminant validity was assessed using the Fornell-Larcker criterion and the Heterotrait–Monotrait (**HTMT**) ratio. As shown in **Table 3**, the square roots of the **AVEs** (diagonal values) are greater than the inter-construct correlations, satisfying the Fornell-Larcker criterion. Additionally, all **HTMT** values are below the conservative threshold of 0.85 (Henseler et al., 2015), confirming discriminant validity.

**Table 3. CAITIZEN measurement model discriminant validity**

Fornell & Larcker Criteria (Diagonal= Root Square -AVE-) for discriminant validity						
HTMT Criteria Ratio <= 0.85 <= 0.90 for convergent validity						
Construct	AFDJ	CAIL	CAITIZEN	EAR	HAIC	MTPP
<b>AFDJ</b>	0.762	<b>0.794</b>	<b>0.824</b>	<b>0.827</b>	<b>0.788</b>	<b>0.763</b>
<b>CAIL</b>	0.718	0.722	<b>0.679</b>	<b>0.834</b>	<b>0.694</b>	<b>0.732</b>
<b>CAITIZEN</b>	0.759	0.614	0.743	<b>0.712</b>	<b>0.820</b>	<b>0.744</b>
<b>EAR</b>	0.766	0.759	0.661	0.745	<b>0.712</b>	<b>0.716</b>
<b>HAIC</b>	0.735	0.637	0.765	0.673	0.810	<b>0.752</b>
<b>MTPP</b>	0.708	0.665	0.695	0.669	0.713	0.760

**Notes:**

**HTMT**. The Heterotrait–Monotrait ratio of correlations assesses discriminant validity by evaluating whether constructs are empirically distinct. Values below 0.85 indicate adequate discriminant validity under a conservative criterion, whereas values below 0.90 may be acceptable for conceptually related constructs (Henseler et al., 2015). Bootstrapping can be used to test whether HTMT confidence intervals include the value of 1.0, providing additional statistical evidence for discriminant validity. Franke and Sarstedt (2019) compare HTMT with other discriminant validity procedures and support its use as a more reliable approach than traditional criteria.

Source: Own elaboration using SmartPLS 4.1.1.8

### 5.1.3. Collinearity Assessment

Before assessing the structural paths, collinearity among predictor constructs was examined using the **inner variance inflation factor (VIF)** values obtained from the **PLS-SEM Algorithm** in **SmartPLS**. The results ranged from **1.000** to **3.269**, remaining below the commonly accepted threshold of **5.0** and also below the more conservative reference value of **3.3**. Therefore, collinearity did not represent a critical issue in the structural model, and the path coefficients could be interpreted without evidence of problematic predictor redundancy. See **Table 4**.

**Table 4. CAITIZEN collinearity assessment of the CAITIZEN structural model**

Structural path	Inner VIF	Assessment
<b>AFDJ -&gt; CAITIZEN</b>	3.269	Acceptable
<b>CAIL -&gt; AFDJ</b>	1.000	Acceptable
<b>CAIL -&gt; EAR</b>	1.000	Acceptable
<b>CAIL -&gt; HAIC</b>	1.000	Acceptable
<b>CAIL -&gt; MTPP</b>	1.000	Acceptable
<b>EAR -&gt; CAITIZEN</b>	2.663	Acceptable
<b>HAIC -&gt; CAITIZEN</b>	2.664	Acceptable
<b>MTPP -&gt; CAITIZEN</b>	2.485	Acceptable

**Note.** VIF = variance inflation factor. Values below 5.0 indicate absence of critical collinearity; values below 3.3 also satisfy a more conservative reference criterion (Hair et al., 2022).

Source: Own elaboration using SmartPLS 4.1.1.8.

### 5.2 Structural Model Assessment

To evaluate the structural model of the **CAITIZEN** framework, the recommended **PLS-SEM** criteria were applied (Hair et al., 2022). First, multicollinearity was examined using the inner Variance Inflation Factor (**VIF**). All **VIF** values were below the conservative threshold of 3.3, indicating no critical collinearity among the predictor constructs and supporting the stability of the path estimates.

After confirming the absence of collinearity problems, the structural relationships were assessed through path coefficients, bootstrapping, confidence intervals, significance levels, and effect sizes. The results generated with **SmartPLS 4.1.1.8** are presented in **Figure 2** and **Table 5**. These results show the empirically validated relationships of the extended **CAITIZEN** framework, in which **CAIL** functions as the foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**, while **AFDJ**, **HAIC**, and **MTPP** significantly predict **AI-assisted Sustainable Citizenship**. The path from **EAR** to **CAITIZEN** was retained in the model but was not statistically significant. This structural assessment supports the explanatory capacity of the model and confirms that **AI-assisted**

sustainable citizenship is mainly shaped by applied capacities related to fairness awareness, human–AI creative collaboration, and metacognitive prompting practices. See Table 5.

**Table 5. Structural model assessment of the CAITIZEN framework: hypothesis testing and effect sizes**

Hypotheses	β Paths	* β Path [t-value; p-value]	Result	*Interval	f <sup>2</sup> Effect Size (pls- algorithm)	
				95% CI [2.5%; 97.5%]	(0.02<=; 0.15 <=0.35)	Effect (Small; Medium; Large)
<b>H1a:</b> “CAIL positively and significantly predicts EAR.”	CAIL -> EAR	0.760 [28.411; 0.000]	Accepted	[0.700; 0.807]	1.371	Large
<b>H1b:</b> “CAIL positively and significantly predicts AFDJ.”	CAIL -> AFDJ	0.720 [24.144; 0.000]	Accepted	[0.655; 0.774]	1.079	Large
<b>H1c:</b> “CAIL positively and significantly predicts HAIC.”	CAIL -> HAIC	0.640 [17.726; 0.000]	Accepted	[0.560; 0.703]	0.692	Large
<b>H1d:</b> “CAIL positively and significantly predicts MTPP.”	CAIL -> MTPP	0.664 [20.186; 0.000]	Accepted	[0.593; 0.722]	0.789	Large
<b>H2:</b> “EAR positively and significantly predicts CAITIZEN.”	EAR -> CAITIZEN	0.045 [0.880; 0.379]	Rejected	[-0.056; 0.144]	0.002	Small
<b>H3:</b> “AFDJ positively and significantly predicts CAITIZEN.”	AFDJ -> CAITIZEN	0.333 [5.498; 0.000]	Accepted	[0.215; 0.450]	0.106	Small
<b>H4:</b> “HAIC positively and significantly predicts CAITIZEN.”	HAIC -> CAITIZEN	0.376 [7.071; 0.000]	Accepted	[0.269; 0.479]	0.166	Medium
<b>H5:</b> “MTPP positively and significantly predicts CAITIZEN.”	MTPP -> CAITIZEN	0.159 [3.516; 0.000]	Accepted	[0.070; 0.249]	0.032	Small

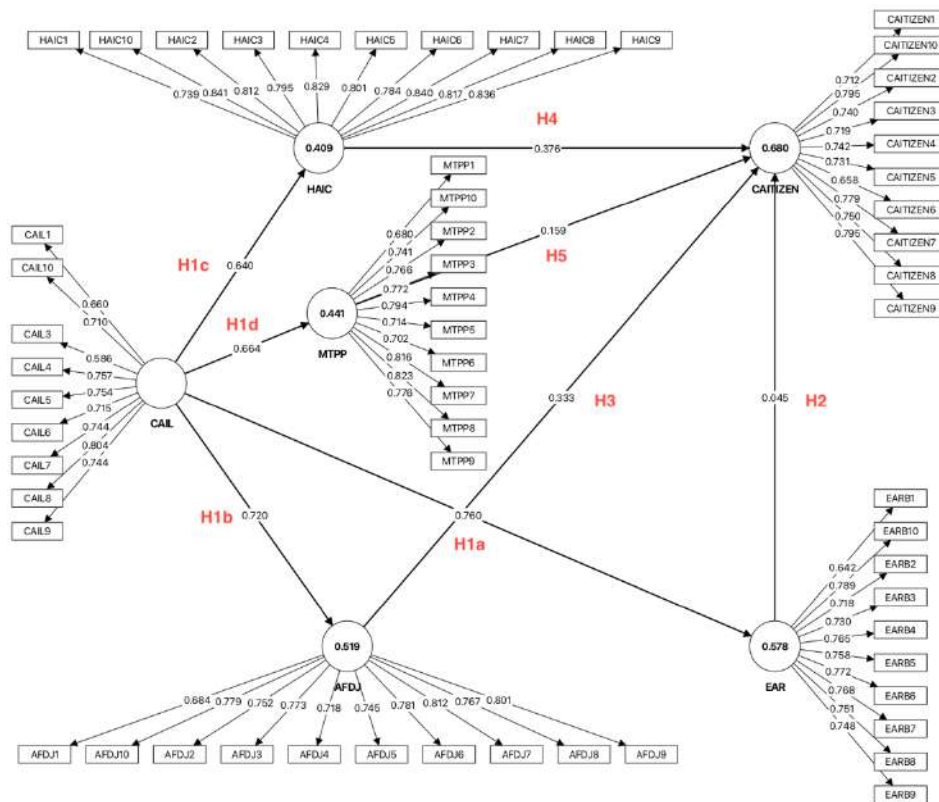
**Notes:**

- **CAIL.** Critical Artificial Intelligence Literacy; **EAR.** Ethical Awareness and Responsibility; **AFDJ.** Awareness of Fairness and Data Justice; **HAIC.** Human–AI Creative Collaboration; **MTPP.** Metacognitive Transparency in Prompting Practices; **CAITIZEN.** AI-assisted Sustainable Citizenship.
- Bootstrapping results are reported as two-tailed t-values and p-values using bias-corrected 95% confidence intervals [2.5%; 97.5%], based on 5,000 subsamples and a 0.05 significance level.
- f<sup>2</sup> effect size. 0.02, 0.15, and 0.35 are interpreted as small, medium, and large (Cohen, 1992)
- Hypothesis acceptance or rejection was determined based on the significance of the bootstrapped path coefficients, whereas f<sup>2</sup> values were used only to interpret effect size magnitude.

Source: Own elaboration using SmartPLS 4.1.1.8

The empirical structural model is presented in **Figure 2** to visually summarize the validated relationships of the extended **CAITIZEN** framework. The figure shows **CAIL** as the foundational antecedent of **EAR**, **AFDJ**, **HAIC**, and **MTPP**. The path from **EAR** to **CAITIZEN** is retained in the model but was not statistically significant. The coefficients displayed in the figure indicate the strength and direction of each structural relationship, whereas the values inside the endogenous constructs represent their explained variance. This visualization complements the hypothesis testing results by showing which relationships were empirically supported and how the constructs interact within the extended model. The strongest effects are observed from **CAIL** toward the proximal capacities, confirming its role as the structural literacy foundation of the model. In contrast, **CAITIZEN** is mainly shaped by applied and actionable dimensions, particularly **HAIC**, **AFDJ**, and **MTPP** suggesting that sustainable AI-assisted citizenship depends not only on ethical awareness, but also on creative agency, fairness-oriented judgment, and reflective prompting practices.

**Figure 2. The CAITIZEN ex-post PLS-SEM structural model results with validated paths and effect sizes**



Source: Own elaboration using SmartPLS 4.1.1.8.

### 5.3. Predictive Relevance

Predictive relevance was assessed using **PLSpredict** in **SmartPLS 4.1.1.8**, focusing on the ten indicators of **CAITIZEN**, understood as **Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked Formation**. **Q<sup>2</sup>\_predict** values greater than zero indicate predictive relevance, while lower **RMSE** values indicate better predictive performance. The **PLS-SEM** model was compared with the linear model benchmark (**LM**). See **Table 6**.

**Table 6. PLSpredict assessment for CAITIZEN**

Indicator	Q <sup>2</sup> predict	PLS-SEM RMSE	LM RMSE	PLS-LM RMSE	Assessment
CAITIZEN1	0.217	1.353	1.370	-0.017	PLS better
CAITIZEN2	0.186	1.226	1.249	-0.023	PLS better
CAITIZEN3	0.157	1.305	1.324	-0.019	PLS better
CAITIZEN4	0.151	1.383	1.394	-0.011	PLS better
CAITIZEN5	0.185	1.299	1.321	-0.022	PLS better
CAITIZEN6	0.150	1.462	1.475	-0.013	PLS better
CAITIZEN7	0.215	1.307	1.331	-0.024	PLS better
CAITIZEN8	0.214	1.220	1.237	-0.017	PLS better
CAITIZEN9	0.247	1.178	1.192	-0.014	PLS better
CAITIZEN10	0.309	1.136	1.149	-0.013	PLS better

**Note. CAITIZEN:** Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked Formation. **Q<sup>2</sup>\_predict** > 0 indicates predictive relevance. **RMSE** = root mean squared error; **LM** = linear model benchmark. Negative **PLS-LM RMSE** values indicate that **PLS-SEM** produced lower prediction errors than **LM**. Source: Own elaboration using SmartPLS 4.1.1.8.

As shown in **Table 6**, all **Q<sup>2</sup>\_predict** values were positive and all **PLS-LM RMSE** differences were negative. Therefore, the extended **CAITIZEN** model demonstrates high out-of-sample predictive power for **Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked Formation**.

## 6. Discussion

The results provide empirical support for the **extended CAITIZEN model** as an **explanatory-predictive framework** for **AI-assisted sustainable citizenship in higher education**. The findings confirm that **Critical Artificial Intelligence Literacy (CAIL)** operates as the **foundational antecedent** of **Ethical Awareness and Responsibility (EAR)**, **Awareness of Fairness and Data Justice (AFDJ)**, **Human–AI Creative Collaboration (HAIC)**, and **Metacognitive Transparency in Prompting Practices (MTPP)**. The strongest structural effects are observed from **CAIL** toward the four proximal capacities, indicating that students’ ability to **understand**,

**evaluate, verify, and responsibly use AI systems** enables more specific forms of **ethical, fairness-oriented, creative, and metacognitive engagement**.

The interpretation of **effect sizes ( $f^2$ )** further clarifies the relevance of the structural results. The effects of **CAIL on EAR ( $f^2 = 1.371$ )**, **AFDJ ( $f^2 = 1.079$ )**, **HAIC ( $f^2 = 0.692$ )**, and **MTPP ( $f^2 = 0.789$ )** are **large**, confirming that **critical AI literacy is the strongest formative foundation of the model**. At the CAITIZEN level, **HAIC shows a medium effect ( $f^2 = 0.166$ )**, making **human–AI creative collaboration** the strongest direct actionable mechanism for forming AI-assisted sustainable citizenship. **AFDJ ( $f^2 = 0.106$ )** and **MTPP ( $f^2 = 0.032$ )** show **small but meaningful effects**, suggesting that fairness awareness and prompt transparency contribute directly, although more moderately. In contrast, **EAR has a negligible effect ( $f^2 = 0.002$ )**, reinforcing that ethical awareness alone does not directly translate into sustainable AI-assisted citizenship unless it becomes actionable through fairness, creativity, and metacognitive practices.

The predictive relevance results strengthen this interpretation. **All  $Q^2_{\text{predict}}$  values for CAITIZEN indicators are positive**, and **all PLS-LM RMSE differences favor the PLS-SEM model**. Therefore, the model does not only explain relationships among constructs; it also shows **predictive capacity** for the CAITIZEN indicators. Overall, the findings indicate that **AI-assisted sustainable citizenship is not produced by technical AI use alone**. It emerges from the interaction between **critical literacy** and **applied capacities** that connect **fairness, creativity, metacognition, responsibility, and sustainable innovation** in higher education.

### **6.1. Theoretical Contribution**

The extended CAITIZEN model contributes theoretically by transforming a **qualitative ethical–cognitive–social framework** into an **empirically validated explanatory-predictive structure**. The original CAITIZEN framework identified **Critical Artificial Intelligence Literacy, Ethical Awareness and Responsibility, Awareness of Fairness and Data Justice, Human–AI Creative Collaboration, and Metacognitive Transparency in Prompting Practices** as relevant dimensions of **AI-assisted sustainable citizenship**. The present study advances that conceptual foundation by specifying and testing the **internal structural logic** among those dimensions. In doing so, it moves CAITIZEN from **conceptual identification** to **empirical explanation**.

The first theoretical contribution is the positioning of **CAIL as a foundational literacy layer**. The large effects of **CAIL on EAR, AFDJ, HAIC, and MTPP** show that critical AI literacy is not merely one isolated component of AI citizenship; rather, it functions as the **cognitive and evaluative condition** that enables other capacities. Students who understand AI systems, evaluate their outputs, recognize their limitations, and verify information are better positioned to develop **ethical responsibility, fairness awareness, creative collaboration, and metacognitive transparency**. Thus, the model theorizes AI-assisted citizenship as a **literacy-enabled process**, not as a simple consequence of AI access or use.

The second contribution concerns the distinction between **normative awareness** and **actionable capacity**. **EAR is significantly enabled by CAIL, but it does not directly predict CAITIZEN** and shows a negligible direct effect. This result refines responsible AI theory by suggesting that **ethical awareness may operate as a necessary orientation but not as a sufficient direct mechanism**. Sustainable AI-assisted citizenship appears to depend more strongly on capacities that can be enacted in practice. The medium effect of **HAIC** indicates that **creative human–AI collaboration** is a central theoretical mechanism, while the small effects of **AFDJ** and **MTPP** confirm that fairness judgment and reflective prompting also contribute to the formation of **CAITIZEN**.

The third contribution is the integration of **multidisciplinarity into a measurable structural model**. **CAITIZEN** articulates **education, artificial intelligence, ethics, data justice, creativity, metacognition, innovation, and sustainable development**. The model therefore expands AI literacy theory by connecting **individual competence** with **civic formation** and **sustainable innovation**. It also contributes to innovation studies by framing **AI-assisted citizenship as a social and conceptual innovation** aligned with **quality education** and **institutional transformation**. Finally, the positive predictive results support the **theoretical robustness of CAITIZEN** as a measurable construct for studying **sustainable AI-mediated citizenship in higher education**.

## 6.2. Practical Contribution

The extended **CAITIZEN** model offers practical value for **universities, educators, curriculum designers, institutional leaders, and policymakers** seeking to guide **responsible AI adoption in higher education**. The results indicate that AI education should begin with **Critical Artificial**

**Intelligence Literacy**, because **CAIL** shows **large effects** on the ethical, fairness-oriented, creative, and metacognitive capacities required for sustainable citizenship. Therefore, institutional AI strategies should not be limited to **technical training, productivity improvement, or tool adoption**. They should include systematic formation in **how AI systems work, how outputs should be verified, how bias may emerge, how privacy and digital rights are protected, and how AI use can be made transparent and accountable**.

The significant effect of **AFDJ on CAITIZEN**, although small, shows that **fairness and data justice must become explicit curricular and institutional priorities**. Universities should train students to identify **algorithmic bias, unequal data representation, opacity, discrimination, and unfair automated decisions**. This implies incorporating **algorithmic fairness modules, case-based learning, data justice discussions, and institutional protocols** for evaluating AI systems used in academic contexts. Even a small direct effect is practically relevant because fairness awareness addresses risks that can affect equity, trust, and responsible participation in AI-mediated environments.

The medium effect of **HAIC on CAITIZEN** shows that AI-assisted citizenship requires **creative agency**. Students should learn to use AI as a **co-creative partner** without surrendering **authorship, judgment, or responsibility**. Educational practice can include activities in which students **compare AI-generated outputs, revise them critically, justify their creative decisions, and document how human agency shaped the final product**. This promotes **responsible innovation** rather than passive dependence on AI systems.

The significant but small effect of **MTPP** highlights the importance of **prompt transparency**. Universities should teach students to **plan, document, monitor, adjust, and evaluate their prompting practices**. **Prompt logs, reflective reports, disclosure statements, and iterative prompt analysis** can become practical tools for **academic integrity and responsible AI use**. Finally, the negligible direct effect of **EAR** suggests that **ethical campaigns and declarations are insufficient if they are not connected to observable practices**. Institutions should translate ethical principles into measurable competencies such as **bias detection, prompt traceability, responsible disclosure, output verification, and human accountability**. In this sense, **CAITIZEN can serve as a diagnostic and formative tool** to assess student readiness, guide

teacher training, design AI policies, and monitor institutional progress toward **responsible AI-assisted sustainable citizenship**.

## 7. CONCLUSION

This section summarizes how the study addresses the research question, the hypotheses, the main findings, and the final scope of the research. The conclusion emphasizes the contribution of the extended **CAITIZEN** model as a quantitative re-specification of the qualitative framework proposed by Mejía-Trejo (2025a), highlighting its value for multidisciplinary, innovation, and sustainable development in higher education.

### 7.1. How the Research Question and Hypotheses Are Addressed

This study addresses the research question by empirically examining how Critical Artificial Intelligence Literacy (**CAIL**), Ethical Awareness and Responsibility (**EAR**), Awareness of Fairness and Data Justice (**AFDJ**), Human–AI Creative Collaboration (**HAIC**), and Metacognitive Transparency in Prompting Practices (**MTPP**) explain **CAITIZEN**, understood as Citizenship Assisted by Artificial Intelligence for Sustainable, Ethical, and Networked Formation.

The results confirm that **CAIL** positively and significantly predicts **EAR**, **AFDJ**, **HAIC**, and **MTPP**; therefore, **H1a**, **H1b**, **H1c**, and **H1d** are accepted. These findings establish **CAIL** as the foundational antecedent of the extended model. In the second structural stage, **AFDJ**, **HAIC**, and **MTPP** positively and significantly predict **CAITIZEN**; therefore, **H3**, **H4**, and **H5** are accepted. However, **EAR** does not significantly predict **CAITIZEN**; therefore, **H2 is rejected**. This result indicates that ethical awareness is relevant, but it requires translation into actionable capacities such as fairness judgment, creative collaboration, and metacognitive prompting.

### 7.2. Research Findings

The main finding is that the extended **CAITIZEN** model advances the original qualitative framework by explaining the internal structural logic among its dimensions. The model does not only identify five components of **AI-assisted sustainable citizenship**; it shows how they interact. **CAIL** operates as the literacy foundation that enables ethical, fairness-oriented, creative, and

metacognitive capacities. In turn, **AFDJ, HAIC, and MTPP** function as proximal mechanisms that directly shape **CAITIZEN**.

The theoretical contribution is that the study provides new knowledge on AI-assisted citizenship as a multidimensional, literacy-enabled, justice-oriented, creative, and reflective construct. This supports the integration of **AI** literacy, responsible **AI**, data justice, human–AI collaboration, metacognition, innovation, and sustainable development. The practical contribution is that the model offers universities a measurable framework for designing **AI** curricula, institutional **AI** policies, responsible prompting practices, fairness education, and innovation-oriented citizenship formation. The positive PLSpredict results further support the model’s out-of-sample predictive relevance.

### **7.3. Final Scope of the Research**

The scope of this research is centered on higher education students who use **AI** tools in academic, creative, and decision-making contexts. The study contributes to **SDG4** by supporting quality education through critical and responsible **AI** competencies, and to **SDG9** by promoting innovation-oriented learning and institutional transformation. However, the results are limited by the use of a non-probabilistic sample, the predominance of undergraduate students, the cross-sectional design, and the absence of experimental or longitudinal evidence.

Future research should test the extended **CAITIZEN** model in different universities, countries, disciplines, and educational levels. Comparative, longitudinal, and mixed-method studies may examine whether the rejected path from **EAR** to **CAITIZEN** becomes significant in other contexts or when mediated by fairness, creativity, or metacognitive practices. Overall, the study confirms that **CAITIZEN** is not only a qualitative conceptual model, but an empirically validated innovation for understanding and forming sustainable **AI**-assisted citizenship in higher education.

The extended **CAITIZEN** model complements the original qualitative model by moving from conceptual substantiation to empirical validation. While the qualitative study established **CAITIZEN** as an ethical–cognitive–social system composed of **CAIL, EAR, AFDJ, HAIC, and MTPP**, the extended model specifies how these dimensions operate structurally. Its main contribution is to demonstrate that **CAIL** functions as the foundational antecedent of the model, whereas **AFDJ, HAIC, and MTPP** act as the direct actionable mechanisms that predict AI-assisted

sustainable citizenship. Thus, the extended model does not replace the original CAITIZEN framework; it advances it by providing explanatory hierarchy, effect-size interpretation, predictive relevance, and practical priorities for higher education institutions

## 8. REFERENCES

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.  
<https://www2.psych.ubc.ca/~schaller/528Readings/Cohen1992.pdf>
- Córdova-Esparza, D.-M. (2025). AI-powered educational agents: Opportunities, innovations, and ethical challenges. *Information*, 16(6), 469. <https://doi.org/10.3390/info16060469>
- Decker, M., Wegner, L., & Leicht-Scholten, C. (2025). Procedural fairness in algorithmic decision-making: The role of public engagement. *Ethics and Information Technology*, 27, Article 1. <https://doi.org/10.1007/s10676-024-09811-4>
- Demirchyan, G. (2025). Algorithmic fairness: Challenges to building an effective regulatory regime. *Frontiers in Artificial Intelligence*, 8, Article 1637134.  
<https://doi.org/10.3389/frai.2025.1637134>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.  
<https://doi.org/10.1177/002224378101800104>
- Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: A comparison of four procedures. *Internet Research*, 29(3), 430–447.  
<https://doi.org/10.1108/IntR-12-2017-0515>
- Georgieva, I., & Georgiev, G. V. (2025). Exploring the use of generative text AI in design creativity inquiries. *Computers in Human Behavior: Artificial Humans*, 6, Article 100219.  
<https://doi.org/10.1016/j.chbah.2025.100219>
- González-Argote, J., Maldonado, E., & Maldonado, K. (2025). Algorithmic bias and data justice: Ethical challenges in artificial intelligence systems. *EthAIca*, 4, Article 159.  
<https://ai.ageditor.ar/index.php/ai/article/view/159>
- Gunasekara, L., El-Haber, N., Nagpal, S., Moraliyage, H., Issadeen, Z., Manic, M., & De Silva, D. (2025). A systematic review of responsible artificial intelligence principles and practice. *Applied System Innovation*, 8(4), 97. <https://doi.org/10.3390/asi8040097>
- Haidar, H., Suryoputro, G., & Safi'i, I. (2025). Impact of the integration of metacognitive prompts by generative artificial intelligence (GenAI) in collaborative and individual learning in improving writing skills and metacognitive awareness. *International Journal of Learning, Teaching and Educational Research*, 24(6), 232–250.  
<https://doi.org/10.26803/ijlter.24.6.11>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.  
<https://us.sagepub.com/en-us/nam/a-primer-on-partial-least-squares-structural-equation-modeling-pls-sem/book270548>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.  
<https://doi.org/10.1108/EBR-11-2018-0203>

- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43, 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- INEGI. (2023). *Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2023*. Instituto Nacional de Estadística y Geografía. <https://www.inegi.org.mx/programas/endutih/2023/>
- Kong, S. C., & Zhu, J. (2025). Developing and validating an artificial intelligence ethical awareness scale for secondary and university students: Cultivating ethical awareness through problem-solving with artificial intelligence tools. *Computers and Education: Artificial Intelligence*, 9, Article 100447. <https://doi.org/10.1016/j.caeai.2025.100447>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- Mejía-Trejo, J. (2025a). *Innovating sustainable artificial intelligence citizenship: A qualitative study of the CAITIZEN model using ATLAS.ti*. *Scientia et PRAXIS*, 5(10), 126–154. <https://doi.org/10.55965/setp.5.10.a5>
- Mejía-Trejo, J. (2025b). *Inteligencia artificial y su repercusión en la educación superior*. AMIDI Editorial. <https://doi.org/10.55965/abib.9786076984543>
- Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- OECD. (2025). *Bridging the AI skills gap: Is training keeping up?* OECD Publishing. [https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/04/bridging-the-aiskillsgap\\_b43c7c4a/66d0702e-en.pdf](https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/04/bridging-the-aiskillsgap_b43c7c4a/66d0702e-en.pdf)
- OECD & European Commission. (2025). *AI literacy framework for primary and secondary education*. <https://learnworkecosystemlibrary.com/initiatives/ai-literacy-framework-for-primary-secondary-education-oecd-ec/>
- OECD & Eurostat. (2005). *Manual de Oslo: Guía para la recogida e interpretación de datos sobre innovación* (3.<sup>a</sup> ed.). OECD Publishing. <https://doi.org/10.1787/9789264065659-es>
- OECD & Eurostat. (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD Publishing. <https://doi.org/10.1787/9789264304604-en>
- Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, 34(2), Article 101885. <https://doi.org/10.1016/j.jsis.2024.101885>
- Pham, N., Pham Ngoc, H., & Nguyen-Duc, A. (2025). Fairness for machine learning software in education: A systematic mapping study. *Journal of Systems and Software*, 219, Article 112244. <https://doi.org/10.1016/j.jss.2024.112244>
- Rafner, J., Zana, B., Bang Hansen, I., Ceh, S., Sherson, J., Benedek, M., & Lebuda, I. (2025). Agency in human-AI collaboration for image generation and creative writing: Preliminary insights from think-aloud protocols. *Creativity Research Journal*, 1–24. <https://doi.org/10.1080/10400419.2025.2587803>

- Salma, Z., Hijón-Neira, R., & Pizarro, C. (2025). Designing co-creative systems: Five paradoxes in human–AI collaboration. *Information*, 16(10), 909. <https://doi.org/10.3390/info16100909>
- Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of market research* (pp. 587–632). Springer. [https://doi.org/10.1007/978-3-319-57413-4\\_15](https://doi.org/10.1007/978-3-319-57413-4_15)
- Southworth, J., Migliaccio, K., Glover, J., Glover, J. N., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence*, 4, Article 100127. <https://doi.org/10.1016/j.caeai.2023.100127>
- Tsakeni, M., Nwafor, S. C., Mosia, M., & Egara, F. O. (2025). Mapping the scaffolding of metacognition and learning by AI tools in STEM classrooms: A bibliometric–systematic review approach (2005–2025). *Journal of Intelligence*, 13(11), 148. <https://doi.org/10.3390/jintelligence13110148>
- United Nations. (2015). *The 17 Sustainable Development Goals*. <https://sdgs.un.org/goals>
- UNESCO & Cámara Nacional de la Industria Electrónica, de Telecomunicaciones y Tecnologías de la Información. (2025, November 4). *UNESCO and CANIETI, with the Microsoft support, implement a model for ethical and responsible artificial intelligence in Mexican companies*. UNESCO. <https://www.unesco.org/en/articles/unesco-and-canieti-microsoft-support-implement-model-ethical-and-responsible-artificial-intelligence>
- Waalder, P. N., Hussain, M., Molchanov, I., Bongo, L. A., & Elvevåg, B. (2025). Prompt engineering an informational chatbot for education on mental health using a multiagent approach for enhanced compliance with prompt instructions: Algorithm development and validation. *JMIR AI*, 4, Article e69820. <https://doi.org/10.2196/69820>
- Wang, C., & Wang, Z. (2025). Investigating L2 writers’ critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing*, 67, Article 101187. <https://doi.org/10.1016/j.jslw.2025.101187>
- Wang, N., Kim, H., Peng, J., & Wang, J. (2025). Exploring creativity in human–AI co-creation: A comparative study across design experience. *Frontiers in Computer Science*, 7, Article 1672735. <https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2025.1672735/full>
- World Economic Forum. (2025). *Future of Jobs Report 2025*. <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>



This is an open access article distributed under the terms of the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)