



## MEDIOS, TÉCNICAS E INSTRUMENTOS EVALUATIVOS: PROPUESTA DE ESTUDIANTES PARA PROFESOR DE MATEMÁTICA Y CHATBOTS

Verónica Parra<sup>1,2</sup>; Ana Rosa Corica<sup>1,2</sup>; Patricia Sureda<sup>1,2</sup>, Silvia Schiaffino<sup>2</sup>, Daniela Godoy<sup>2</sup>

<sup>1</sup>Núcleo de Investigación en Educación Matemática (NIEM), <sup>2</sup>Instituto Superior de Ingeniería de Software de Tandil (ISISTAN). Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA)-Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

Argentina

[vparra@niem.exa.unicen.edu.ar](mailto:vparra@niem.exa.unicen.edu.ar), [acorica@niem.exa.unicen.edu.ar](mailto:acorica@niem.exa.unicen.edu.ar),  
[psureda@niem.exa.unicen.edu.ar](mailto:psureda@niem.exa.unicen.edu.ar), [silvia.schiaffino@isistan.unicen.edu.ar](mailto:silvia.schiaffino@isistan.unicen.edu.ar),  
[daniela.godoy@isistan.unicen.edu.ar](mailto:daniela.godoy@isistan.unicen.edu.ar)

Evaluación en educación matemática.

Secundaria (12 a 18 años).

**Resumen:** Este trabajo analiza las evaluaciones sobre nociones de estadística, específicamente población y muestra, realizadas por tres futuros profesores de matemática y tres chatbots basados en modelos de Inteligencia Artificial (IA) generativa. Se comparan en términos de medios, técnicas e instrumentos evaluativos. Además, se explora la funcionalidad de los chatbots como asistentes para generar distintos tipos de evaluaciones. Los resultados evidencian diferencias significativas entre las evaluaciones, destacando la utilidad de los chatbots como herramientas valiosas, ya que permiten diseñar evaluaciones tanto tradicionales (por ejemplo, pruebas escritas) como no tradicionales (por ejemplo proyectos de investigación).

**Palabras claves:** evaluación, matemática, población y muestra, profesores en formación, chatbots.

### Introducción.

La evaluación en Matemática se suele asociar únicamente a la calificación numérica y verificación de resultados, limitando su función a acreditar un curso o materia y promover a los estudiantes al siguiente año. Este enfoque reduce la evaluación a un instrumento para conocer qué sabe el estudiante, sin proporcionar información sobre cómo lo sabe, por qué lo sabe, qué no sabe o por qué no lo sabe (Prieto, 2008). En la actualidad, diversos investigadores en Didáctica de la Matemática han cuestionado este reduccionismo de la evaluación (Rodríguez y Pochulu, 2023; Sureda et al., 2024; Zavaleta y Flores, 2021, entre otros) planteando nuevas finalidades e incluso, definiciones diferentes de la noción de “evaluar”. Por ejemplo, Rodríguez y Pochulu (2023) sintetizan las críticas y acuerdos más frecuentes realizadas a la evaluación matemática. Respecto a las críticas, aluden a: falta de conexión con el mundo real; enfoque excesivo en la memorización; abuso de



evaluaciones escritas y ausencia de retroalimentación significativa. Respecto a los acuerdos refieren a: promover una evaluación más comprensiva; fomentar la actividad en el aula; promover la retroalimentación significativa y la evaluación personalizada. En líneas con estos cuestionamientos, este trabajo tiene como objetivo analizar evaluaciones diseñadas por tres futuros profesores de matemática y por tres chatbots basados en modelos de Inteligencia Artificial (IA) generativa. El análisis compara los tipos de evaluaciones propuestas sobre población y muestra, y aporta sobre la utilidad de los chatbots como potenciales herramientas para asistir en la creación de diversas formas de evaluación.

### **Marco teórico: medios, técnicas e instrumentos del sistema evaluativo.**

Se adopta como marco conceptual los componentes del sistema evaluativo definidos por Hamodi et al. (2015): medios, técnicas e instrumentos. Los *medios de evaluación* se refieren a las producciones de los estudiantes que el profesor puede recoger, ver y/o escuchar. Pueden adoptar tres formas diferentes: a) escritos, por ejemplo, examen, carpeta, diario de clase, proyecto, trabajo escrito, etc.; b) orales, por ejemplo, comunicación, debate, presentación oral, etc. y c) prácticos, por ejemplo, práctica supervisada, representación, juego de roles, etc. Las *técnicas de evaluación* son las estrategias que el profesor utiliza para recoger información sobre las producciones y evidencias creadas por los estudiantes de los medios. Las técnicas a utilizar son diferentes en función de si los estudiantes participan o no en el proceso de evaluación. Los *instrumentos de evaluación* corresponden a las herramientas que los estudiantes y el profesor emplean para organizar la información recogida a través de una determinada técnica de evaluación. Algunos ejemplos de estos instrumentos pueden ser: diario del profesor, rúbrica, ficha de observación, ficha de seguimiento individual o grupal, fichas de autoevaluación, fichas de evaluación entre iguales, informe de autoevaluación, etc.

### **Metodología.**

Este trabajo analiza las evaluaciones sobre las nociones de población y muestra propuestas por tres futuros profesores de Matemática de secundaria (P1, P2 y P3) y por tres chatbots: ChatGPT, Gemini y Copilot (creativo, equilibrado y preciso). Los estudiantes para profesor diseñaron estas evaluaciones durante un curso de Didáctica de la Matemática del tercer año de su formación en una universidad Argentina. Se consideran las propuestas de los estudiantes que optaron por las nociones de población y muestra. El análisis se realiza sobre el documento escrito entregado al final del curso. En el caso de los chatbots, la tarea de generación de texto recae en los modelos de lenguaje o LLMs en los que se fundamentan (Kumar et al., 2024). ChatGPT utiliza los modelos de la familia GPT, como GPT 3.5 y GPT 4, este último empleado por Copilot. Gemini utiliza el modelo del mismo nombre, sucesor de PaLM-2 usado por Bard. La generación de contenido en el caso de los chatbots responde al siguiente *prompt*: Eres un profesor de matemática de escuela secundaria y tienes que evaluar el tema de Estadística: población y muestra, a estudiantes que van a la escuela secundaria de Argentina. ¿Cuáles serían las posibles maneras de evaluar el tema?



El análisis de los datos (las evaluaciones generadas por los estudiantes y los chatbots) se realizó a partir de las categorías adoptadas del marco teórico: medios, técnicas e instrumentos de evaluación. A su vez, dentro de cada categoría se consideran subcategorías. Estas últimas fueron generadas de forma inductiva. La Tabla 1 sintetiza esta categorización:

**Tabla 1.** Categorías y subcategorías de análisis.

Estudiantes Chatbots	Medios								Técnicas	Instrumentos	
	Escritos				Orales			Práctico			
	Prueba escrita	Proyecto de investigación	Tarea práctica	Evaluación formativa	Presentación en clase	Examen	Debate	Uso de tecnología/Simuladores	Autoevaluación Entre pares	Infografía	Recurso audio/video

Fuente: Elaboración propia.

## Resultados.

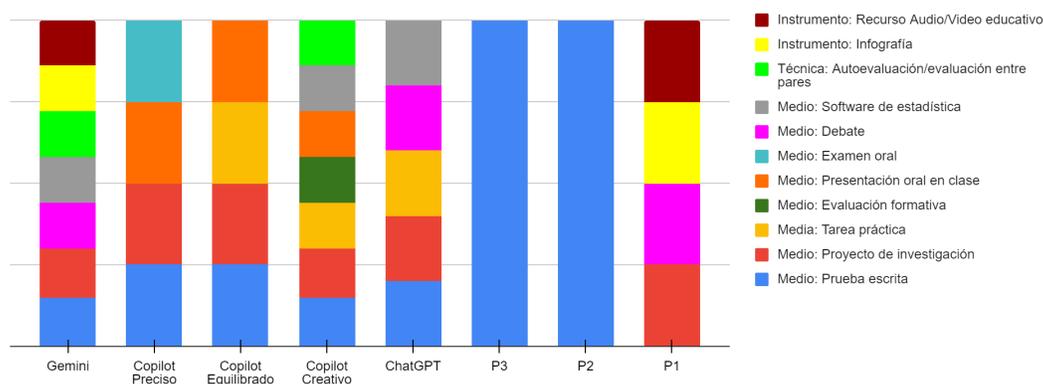
Se identificaron en total (entre estudiantes para profesor y chatbots) ocho medios de evaluación distintos, una única técnica y dos instrumentos. Respecto a los medios, se identificaron cuatro escritos, tres orales y uno práctico. Los medios escritos corresponden a prueba escrita, proyecto de investigación, tarea práctica y evaluación formativa. La prueba escrita se refiere al examen escrito tradicional, compuesto por ejercicios específicos a resolver de manera individual. Este medio de evaluación fue propuesto por todos los chatbots y estudiantes para profesor, excepto por P1. El proyecto de investigación consiste en llevar a cabo una investigación, una exploración o indagación, que involucra los saberes a evaluar, con la presentación de esa investigación en algún formato de entrega, por ejemplo, la redacción de un informe. Este medio fue propuesto por todos los chatbots y sólo por un profesor (P1). La tarea práctica corresponde a la resolución de tareas matemáticas relativas al tema, con datos que brinda el profesor. Este medio de evaluación fue propuesto solamente por Copilot equilibrado, Copilot creativo y ChatGPT. La evaluación formativa se realiza durante todo el proceso de estudio y permite monitorear el progreso de los estudiantes en particular y de la clase en general. No se desarrolla en un momento específico, determinado en tiempo y espacio. Este medio fue propuesto únicamente por Copilot creativo.

Los medios de evaluación de tipo oral identificados son presentación oral en clase, examen oral y debate. El primero se refiere a la presentación del tema a evaluar, mediante algún soporte como puede ser una presentación por diapositivas. Este medio fue propuesto solamente por Copilot (en sus tres modalidades). El examen oral se trata de un tipo de evaluación donde el profesor realiza preguntas a un estudiante por vez, quien debe responder

de manera inmediata, desarrollándose completamente de forma verbal. Fue propuesto por todos los chatbots y los estudiantes, excepto P1. El debate genera un espacio dentro de la clase en el que se realizan discusiones con todos los integrantes del grupo sobre la temática a evaluar, en este caso, población y muestra. Este medio fue propuesto por Gemini, ChatGPT y P1. El medio evaluativo práctico identificado corresponde al uso de tecnología/simuladores, en este caso, un software de estadística. Se trata de un medio de evaluación donde el conocimiento a evaluar se realiza utilizando algún software para analizar un conjunto de datos dados. Este medio fue propuesto por Gemini, Copilot creativo y ChatGPT.

La técnica autoevaluación/evaluación entre pares fue identificada en la propuesta de Gemini y Copilot creativo. Con esta técnica los estudiantes evalúan sus propias producciones y/o de sus compañeros, con el fin de desarrollar habilidades críticas y de auto-reflexión. Los instrumentos de evaluación identificados son infografía y recurso audio/video educativo. La infografía es un instrumento que demanda recuperar lo estudiado, y transformarlo para ser comunicado a través de alguna forma de difusión escrita. Este tipo de instrumento requiere, además del poder de síntesis del emisor, de las habilidades interpretativas del receptor para entender y otorgar sentido a la información comunicada. Este instrumento fue propuesto por Gemini y uno de los estudiantes para profesor (P1). El instrumento recurso audio/video educativo demanda recuperar lo estudiado y transformarlo para ser comunicado a través de algún medio audiovisual de difusión, pero de forma dinámica. De forma análoga a la infografía requiere, además del poder de síntesis del emisor, del orden cronológico y la organización del contenido en una línea temporal. Este instrumento fue también propuesto por Gemini y P1. La figura siguiente sintetiza los resultados.

**Figura 1.** Síntesis de los medios, técnicas e instrumentos propuestos



Fuente: Elaboración propia.

## Conclusiones.

En cuanto a la caracterización y comparación de los medios, técnicas e instrumentos de evaluación propuestos por los futuros profesores de matemática y los chatbots sobre nociones



de estadística (población y muestra), concluimos lo siguiente: Gemini es el único que abarca los tres componentes del sistema evaluativo, con cuatro medios, una técnica y dos instrumentos. Copilot creativo ofrece la mayor variedad de medios (todos excepto examen oral y debate) y una técnica, pero no incluye instrumentos. Los otros chatbots (ChatGPT, Copilot preciso y Copilot equilibrado) se limitan exclusivamente al uso de medios, sin incorporar técnicas ni instrumentos. En el caso de los futuros profesores, P1 considera dos componentes evaluativos distintos: dos medios (proyecto de investigación y debate) y dos instrumentos (infografía y recursos audiovisuales educativos), pero no menciona técnicas. Por su parte, P2 y P3 sólo proponen un medio: la prueba escrita.

Se observa que los chatbots presentan una mayor variedad de medios de evaluación en comparación con los estudiantes para profesor. Una posible explicación radica en la forma en que los chatbots generan sus respuestas, basándose en la amplia y heterogénea información con la que han sido entrenados. Por otro lado, la preferencia de los estudiantes para profesor (dos de tres) por la prueba escrita como único medio de evaluación podría estar influida por restricciones asociadas a la escasez de tiempo, que figura como una de las variables principales de la profesión docente.

### Referencias bibliográficas

- Hamodi, C., López Pastor, V., López Pastor, A. (2015). Medios, técnicas e instrumentos de evaluación formativa y compartida del aprendizaje en educación superior. *Perfiles educativos*, 37(147), 146-161.
- Kumar, P. Large language models (LLMs): survey, technical frameworks, and future challenges. *Artificial Intelligence Review* 57, 260 (2024).
- Prieto, M. (2008). Creencias de los profesores sobre Evaluación y Efectos Incidentales. *Revista de Pedagogía*, 29(84), 123-144.
- Rodriguez, M. y Pochulu, M. (2023). Evaluación de aprendizajes en matemática: perspectivas teóricas y ejemplos. *Unión.*, 67, 1-20.
- Sureda, P., Corica, A., Parra, V., Godoy, D., y Schiaffino, S. La evaluación en educación matemática: aportes de chatbots y futuros profesores de matemática. *Eduotec, Revista Electrónica de Tecnología Educativa*, (89) 64-83.
- Zavaleta, A. y Flores, C.(2021).Evaluación para el aprendizaje en matemáticas: el caso de la retroalimentación. *Números*, 107, 9-34.