

Análisis y comparación de respuestas aportadas por IA generativa y por profesores de matemática a un problema de geometría euclidiana

Analysis and comparison of generative AI and mathematics teachers' answers to an euclidean geometric problem

Ana Rosa Corica,¹ Verónica Parra,² Patricia Sureda,³ Silvia Schiaffino,⁴ Daniela Godoy⁵

Resumen: Este trabajo tiene por objetivo esclarecer el desempeño de la Inteligencia Artificial Generativa (IAG) en la resolución de problemas matemáticos y su impacto en la profesión docente. El marco conceptual adoptado es el Aprendizaje Basado en Problemas. Se exploran, analizan y comparan las

Fecha de recepción: 14 de marzo de 2024. **Fecha de aceptación:** 4 de enero de 2025.

¹ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto Superior de Ingeniería del Software (ISISTAN - CONICET- UNCPBA), Núcleo de Investigación en Educación Matemática (NIEM), Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA) acorica@niem.exa.unicen.edu.ar, <https://orcid.org/0000-0002-3583-6081>

² Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto Superior de Ingeniería del Software (ISISTAN - CONICET- UNCPBA), Núcleo de Investigación en Educación Matemática (NIEM), Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA) vparra@niem.exa.unicen.edu.ar, <https://orcid.org/0000-0002-6956-0052>

³ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto Superior de Ingeniería del Software (ISISTAN - CONICET- UNCPBA), Núcleo de Investigación en Educación Matemática (NIEM), Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA) psureda@niem.exa.unicen.edu.ar, <https://orcid.org/0009-0004-6223-4424>

⁴ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto Superior de Ingeniería del Software (ISISTAN - CONICET- UNCPBA), Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), silvia.schiaffino@isistan.unicen.edu.ar, <https://orcid.org/0000-0002-7616-9756>

⁵ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto Superior de Ingeniería del Software (ISISTAN - CONICET- UNCPBA), Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA) daniela.godoy@isistan.unicen.edu.ar, <https://orcid.org/0000-0002-5185-4570>

respuestas a un problema que ofrecen modelos de IAG y profesores de matemática. El problema es de geometría euclidiana, caracterizado por no admitir respuesta única e inmediata. Se propone dividir un círculo en tres superficies de dimensiones cualesquiera pero con igual área. Las categorías de análisis son: tipo de corte, sistema de representación, nivel de validación matemática y tipo de resolución. Se concluye que la IAG ofrece respuestas inconsistentes, poco variadas, con escasas validaciones matemáticas y cuyo sistema de representación predominante es el coloquial. Las respuestas de los profesores son correctas, ofrecen gran variedad de particiones, preponderan los sistemas de representación geométrico, gráfico y aritmético y ofrecen algún tipo de validación. Se concluye que el empleo de estos modelos de IAG no son apropiados para este tipo de problemas, sin una reflexión crítica.

Palabras clave: *Inteligencia Artificial, Docente de secundaria, Profesor especializado, Geometría.*

Abstract: This work aims to shed some light on the performance of Generative Artificial Intelligence (GAI) in solving mathematical problems and its impact on the teaching profession. The conceptual framework adopted throughout this work is Problem Based Learning. The answers to a problem offered by GAI models and mathematics teachers are explored, analyzed and compared. The problem studied belongs to euclidean geometry, and it is characterized by not admitting a single and immediate answer. It consists of dividing a circle into three surfaces of any dimension but with equal area. The categories for analysis are type of cut, representation system, level of mathematical validation and resolution type. We concluded that GAI offers inconsistent responses, little variety of answers, with few mathematical validations and with a predominance of a colloquial representation system. The teachers' answers are correct, they offer a great variety of partitions, the predominant representation systems are geometric, graphic and arithmetic ones, and they offer some type of validation. It is concluded that the use of GAI models are not appropriate for this type of problems, without any critical reasoning.

Keywords: *Artificial Intelligence, Secondary School Teacher, Specialized Teacher, Geometry.*

1. INTRODUCCIÓN

Los *chatbots* tales como el ChatGPT o Gemini son modelos de Inteligencia Artificial Generativa (IAG) diseñados para generar texto similar al humano. Su irrupción en la sociedad, a finales del año 2022, y la rápida incorporación en el sistema educativo, de la mano de los estudiantes más que de los profesores, generó preocupación, riesgos e incertidumbre en los sistemas escolares (Lozada Lozada *et al.*, 2023; Sanabria-Navarro *et al.*, 2023). Al contrario de las tecnologías previas, como los buscadores, las capacidades de la IAG amenazan la *expertise*, es decir el conocimiento, las habilidades y la experiencia que una persona pueda tener en una determinada área (Romo-Pérez *et al.*, 2023). En particular, estos temores alcanzaron a la comunidad docente, donde incluso se puso en debate el reemplazo de los profesores por la IAG (Gallent-Torres *et al.*, 2023).

Considerando particularmente el temor al reemplazo de los profesores por los recursos de IAG, este trabajo se orienta a llevar tranquilidad en este sentido. Centrándonos en la matemática, nos proponemos comparar las resoluciones de los chatbots y profesores de matemática con el fin de determinar las fortalezas y debilidades de cada uno en la resolución de un problema de geometría clásica, no sólo en términos de lo adecuado de las respuestas, sino también en la variedad en las soluciones aportadas, del sistema de representación de la resolución más preponderante, del nivel de validación matemática y de la representación gráfica de la solución. Pretendemos esclarecer el rol de la IA generativa y delimitar las expectativas respecto de lo que esta tecnología puede aportar a la resolución de problemas matemáticos que podrían ser implementados en el aula.

El problema considerado, denominado problema de la pizza (Camacho Machin *et al.*, 2015), involucra para su resolución nociones de geometría euclidiana. La característica diferencial de este problema, motivo de su elección, es que su solución no es inmediata, ni única. El problema consiste en dividir un círculo (en el enunciado del problema, una pizza) en tres superficies de igual área. Matemáticamente, las maneras de hacer esto son variadas (más aún, infinitas), y dependiendo de la partición propuesta, la matemática necesaria para resolverlo puede ser hasta cierto punto compleja. Además de las diferentes opciones del corte, el problema adiciona la característica propia de una validación. Es decir, se propone un posible corte, en el que resultan 3 superficies (que no necesariamente tienen igual dimensión), y se debe demostrar que corresponden a figuras de igual área. Esta variabilidad permite también que el problema de la pizza pueda ser abordado en diferentes años del nivel secundario. La impronta

de los diseños curriculares oficiales de matemática en Argentina es “enseñar matemática a partir de la resolución de problemas”. Por ejemplo, se enfatiza que “Hacer Matemática es, básicamente, resolver problemas; por lo tanto esta tarea deberá ocupar un lugar central en su enseñanza” (Dirección General de Escuelas de la Provincia de Buenos Aires, 2006, p. 295).

En la siguiente sección se describen antecedentes relevantes respecto a la IAG y Educación Matemática. En la sección 3 se describe el referente conceptual adoptado: Aprendizaje Basado en Problemas (ABP). En la sección 4, se caracteriza el problema de la pizza, su origen, enunciado y algunas posibles soluciones del mismo. En la sección siguiente, se presenta la metodología de la investigación. Finalmente, en las secciones 6 y 7, los resultados y su discusión y las conclusiones.

2. ANTECEDENTES: IAG Y EDUCACIÓN MATEMÁTICA

Alrededor del año 2000, el uso de las Tecnologías de la Información y la Comunicación (TIC), en el ámbito educativo experimentó un cambio significativo y revolucionario con la incorporación de herramientas que facilitaron la interacción entre las personas mediadas por tecnología (Peralta Roncal *et al.*, 2023). Dentro de este campo, diversas investigaciones abordan el uso de las TIC en la enseñanza y aprendizaje de la matemática (Burgos Cantos y Cao, 2023; Grisales-Aguirre, 2018; Vaillant *et al.*, 2020). Sin embargo, los recursos provenientes de la IAG poseen características distintivas, como son la facilidad de interacción en lenguaje natural y la posibilidad de brindar respuestas verosímiles a las preguntas planteadas a casi cualquier tema sobre los que sean consultados. Selwyn *et al.* (2022) destacan que la IAG es una tecnología poderosa y el desafío se ubica en poder establecer las formas de uso con sentido y con criterio. La IAG ha sido desarrollada con el objeto de generar nuevo contenido, ya sea textual o multimodal. Los *chatbots* son uno de los productos más difundidos de la IAG, orientados a mantener una conversación fluida y amigable con los usuarios a partir del uso de grandes modelos de lenguaje o LLMs (del inglés *Large Language Models*). Estos son modelos creados con técnicas de aprendizaje profundo a partir de grandes volúmenes de texto que capturan las relaciones entre las palabras y los patrones del lenguaje, considerando incluso las sutilezas y matices en su uso.

Si bien existen algunos trabajos que abordan el uso de la IAG en educación matemática (Silva *et al.*, 2024; Stefanova y Georgiev, 2024), son pocos los que

comparan el desempeño de profesores de matemática en servicio con el desempeño de los chatbots ante un mismo problema matemático (Corica et al 2024; Sureda et al., 2024). En Sureda et al. (2024) se analizan las evaluaciones propuestas por futuros profesores de matemática y por chatbots basados en IAG, comparando los tipos de evaluaciones propuestas sobre nociones de estadística (población y muestra) y se determina la funcionalidad de los chatbots como posibles asistentes para la generación de diferentes tipos de evaluaciones. Se concluye que los chatbots pueden resultar en asistentes valiosos a la hora de crear evaluaciones, ya que ofrecen diferentes tipos, tanto tradicionales, por ejemplo una prueba escrita, como no tradicionales, como un proyecto de investigación. Stefanova y Georgiev (2024) sostienen que la IA no reemplazará a los educadores, pero puede usarse para liberar parte de su tiempo.

3. MARCO CONCEPTUAL: EL APRENDIZAJE BASADO EN PROBLEMAS (ABP)

La noción de problema adquiere diversos significados según la perspectiva teórica que se adopte. Polya (1981) lo define del siguiente modo: “Tener un problema significa buscar de forma consciente una acción apropiada para lograr un objetivo claramente concebido, pero no alcanzable de forma inmediata” (p.117). Este autor junto a otros como Schoenfeld (1992), dieron lugar al enfoque de Aprendizaje Basado en Problemas (ABP). El ABP se ha constituido desde entonces como una estrategia didáctica y metodológica transversal a las diferentes disciplinas (Padilla-Doria y Flórez-Nisperuza, 2022). La esencia del ABP consiste en identificar, describir, analizar y resolver problemas de forma colaborativa entre profesores y estudiantes (Velázquez et al., 2021).

Diversos autores exploraron la metodología del ABP en distintos niveles educativos. En el nivel secundario, Espinoza Melo y Sánchez Soto (2014) emplearon el ABP para enseñar nociones de probabilidad y estadística a estudiantes de cuarto año de educación media (17-18 años de edad) en Chile. Los autores lograron integrar el ABP al aprendizaje significativo resultando en una valoración favorable de esa integración. García (2016) realizó una propuesta basada en ABP para estudiantes de secundaria (14-15 años de edad) con el objetivo de enseñar nociones geométricas tales como perímetro y área de polígonos y figuras circulares, triángulos semejantes, medidas de longitud y superficie, área y volumen de poliedros cilindros, conos y esferas y la identificación de ciertos elementos tales como centro, ejes y planos de simetría de geometría. Los

resultados indican que los estudiantes se involucraron mucho más en el aprendizaje de la geometría e incluso, en las instancias de evaluación.

En la formación de profesores de matemática, Molina-Patiño *et al.* (2023) indagan, a través de un cuestionario, sobre el conocimiento de la metodología ABP que tiene un grupo de 33 futuros profesores de matemática, en la formación de sus prácticas profesionales. Los resultados indican que, a pesar de la importancia atribuida a una enseñanza por ABP, se identificaba en estos futuros profesores, rasgos de una enseñanza tradicional. Espinoza-Benavidez y Triminio-Zavala (2018) realizan entrevistas semiestructuradas a cinco docentes con el objetivo de conocer la implementación de la metodología ABP para enseñar geometría plana. Los autores concluyen que el uso de ABP en geometría plana fue la forma más eficaz de mejorar los espacios de aprendizaje que actualmente se desarrollan en sus aulas.

En este trabajo se comparan las resoluciones que ofrecen tres chatbots (ChatGPT, Gemini y Copilot) con las que realizan ocho profesores de matemática a un problema de geometría clásica. Esta comparación no sólo se realiza en términos de si las respuestas son correctas o incorrectas, sino también en función de la variedad de las soluciones aportadas, del sistema de representación de la resolución más preponderante y del nivel de validación matemática. De esta forma, se busca aclarar el rol y la incidencia de la IAG en el bagaje de recursos de los profesores de matemática delimitando las expectativas respecto de lo que esta tecnología puede aportar a la resolución de problemas de matemática. En la sección siguiente se describe el problema y algunas de las posibles formas de resolverlo. Conviene destacar que por cuestiones de espacio no es posible precisar todas las maneras de resolver el problema y menos aún detallar las justificaciones matemáticas asociadas a cada respuesta, pero consideramos que un lector con una formación en Matemática puede profundizar las solución a partir de las soluciones parciales que ofrecemos en el manuscrito.

4. EL PROBLEMA DE LA PIZZA

La elección de este problema, en particular, se fundamenta en distintos factores. En primer lugar, se alinea con la impronta de los diseños curriculares de la Provincia de Buenos Aires (Argentina), que es la enseñanza de la matemática a través de la resolución de problemas. En segundo lugar, por tratarse de un problema de geometría, se busca recuperar un área de la matemática poco

abordada en las aulas, más allá de su integración en los diseños curriculares (Itzcovich, 2005; Rojas y Sierra, 2020). Itzcovich (2005) resalta la dificultad de los profesores en encontrar problemas que representen verdaderos desafíos para los estudiantes. Así, en tercer lugar, por tratarse de un problema que no es de respuesta inmediata y que posee múltiples soluciones, se trata de un problema con potencialidad para que los profesores lo implementen en sus aulas. Así, en función de la resolución escogida, el problema se puede adaptar a distintos niveles de la escolaridad secundaria. Por ejemplo, un corte simple permitiría abordar nociones básicas de círculo y de circunferencia (ángulo central, sector circular, segmento circular, etc.), lo cual se corresponde a un ciclo básico de secundaria. Mientras que un corte más complejo, involucraría nociones matemáticas del nivel superior (faja central, faja lateral, circunferencia goniométrica, etc.). Por último, el problema no requiere formación adicional del profesor de matemática, por lo que puede ser llevado con facilidad al aula.

4.1. ORIGEN Y ENUNCIADO DEL PROBLEMA

El problema de la pizza se deriva de un teorema denominado, análogamente, *El Teorema de la pizza*, publicado por Upton en 1967, como un problema a resolver:

Four lines in a plane are concurrent at O . The angles between the lines are each 45° . A circle is superimposed on this configuration so that O lies within the circle. (a) Show that the alternate sectors cover one-half of the circle. (b) Show this result without use of the calculus. (Kravitz *et al.*, 1967, p. 3)⁶

Esta idea del problema deriva en el Teorema de la Pizza, donde se propone particionar el círculo de forma análoga a la propuesta en el problema, con el punto de intersección de las rectas desplazadas del centro. Desde la demostración realizada por Nelsen (1993) hasta la actualidad, este problema está presente en la comunidad matemática, generando diversas publicaciones con diferentes posibles cortes. La más reciente corresponde al método denominado *mosaico de disco monoédrico* (Haddley y Worsley, 2015).

⁶ Cuatro líneas en un plano son concurrentes en O . Los ángulos entre las líneas son cada uno de 45° . Se superpone un círculo a esta configuración de modo que O se encuentre dentro del círculo. (a) Demuestre que los sectores alternos cubren la mitad del círculo. (b) Demuestre este resultado sin utilizar el cálculo.

4.2. ALGUNAS POSIBLES SOLUCIONES DEL PROBLEMA

El problema, tal como fue propuesto a los profesores y a los chatbots, tiene el siguiente enunciado:

La pizza es un alimento que forma parte de nuestra dieta y que todos sabemos partir. El corte tradicional en forma de sector circular es el más extendido, pero ¿es la única forma de dividir una pizza en tres partes iguales? (Camacho Machín *et al.*, 2015).

Para su resolución, cualquiera sea el corte propuesto, es necesario detallar la forma de hacerlo, la construcción geométrica del mismo y validar que las tres superficies generadas tengan igual área. Esto se realiza con conocimientos y procedimientos geométricos de diferentes niveles de complejidad dependiendo del tipo de corte propuesto. Por ejemplo, algunos posibles cortes son los que se describen en las siguientes subsecciones.

4.2.1. División de la pizza en trozos tradicionales: sectores circulares

Este corte es el tradicional: cada una de las porciones están delimitadas por dos radios R , un arco de circunferencia C y un ángulo α , que es el ángulo que hay entre los dos radios del sector (amplitud del ángulo central del sector).

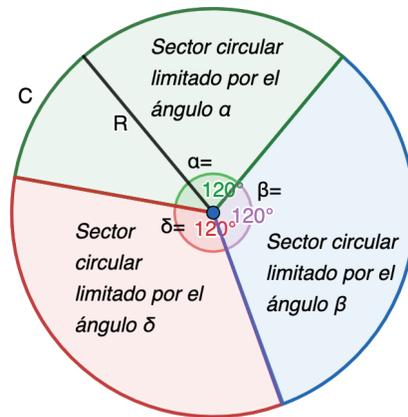


Figura 1. Sectores circulares (elaboración propia).

Se demuestra que los tres sectores circulares conforman figuras de igual área pues los ángulos α , β y γ tienen la misma amplitud, 120° . El área de cada sector circular es $S = \frac{\pi R^2 \cdot \alpha}{360^\circ} = \frac{\pi R^2 \cdot 120^\circ}{360^\circ} = \frac{1}{3} \pi R^2$.

4.2.2. División de la pizza en trozos concéntricos: coronas circulares

Este corte corresponde a considerar dos coronas circulares y un círculo central. A partir de la circunferencia C de centro O y radio R , se trazan dos circunferencias concéntricas, denotadas por C_1 y C_2 . Los radios de cada una de estas circunferencias, respectivamente, son denotados por r_1 y r_2 .

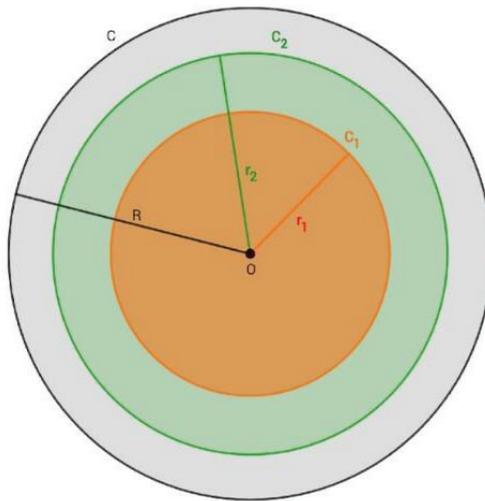


Figura 2. Coronas circulares y círculo concéntricos (elaboración propia).

La pizza se puede particionar en tres superficies de igual área, donde cada parte corresponde al círculo determinado por la circunferencia C_1 (zona anaranjada en la figura 2), la corona circular denominada *central*, delimitada por las circunferencias C_2 y C_1 (zona verdosa en la figura 2), y la corona circular delimitada por las circunferencias C y C_2 (zona gris en la figura 2). Para determinar tres superficies de igual área se debe establecer la relación entre los tres radios: R , r_1 y r_2 . Se concluye que $r_1 = \frac{\sqrt{3}}{3} R$ y $r_2 = \frac{\sqrt{6}}{3} R$ lo que permite demostrar la equivalencia de áreas a partir de estos radios.

4.2.3. Divisiones de la pizza en forma de lágrima: trisección del radio (y diámetro)

La clave de este corte es la trisección del radio R de la circunferencia C de centro O . Esta trisección puede realizarse con diversos procedimientos matemáticos que involucran nociones tales como Teorema de Thales, vectores, entre otros. Luego, formar semicírculos cuyos centros son puntos correspondientes a esa trisección. Considerando estos semicírculos, se forman tres superficies denominadas *lágrimas*. Por ejemplo, para la lágrima de la izquierda (de color anaranjado en la figura 3), se considera el semicírculo inferior de centro A_1 y radio A_1A_2 más la superficie resultante de quitarle al semicírculo superior de centro O y radio R el semicírculo superior de centro A_3 y radio A_3B . De forma análoga, se construyen las dos restantes lágrimas: la central (de color verde en la figura 3) y la derecha (de color gris en la figura 3).

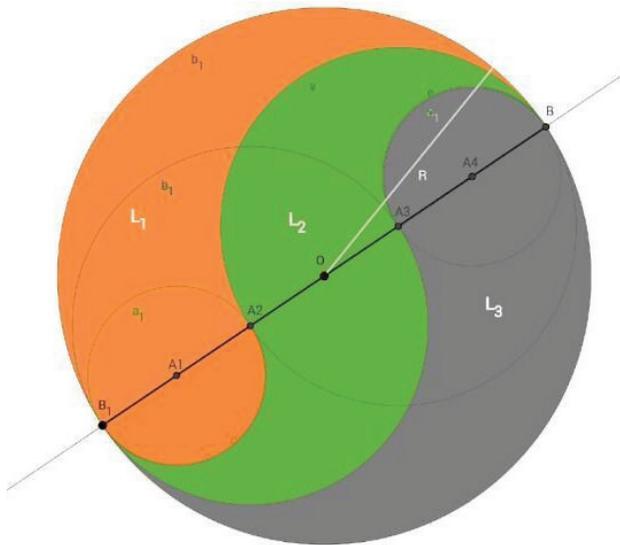


Figura 3. Formas de lágrimas (elaboración propia).

Para que la pizza se particione en tres partes equivalentes, se debe demostrar que las tres lágrimas son figuras de igual área.

4.2.4. Divisiones de la pizza usando rectas paralelas: segmentos circulares y faja central

Este corte implica nociones tales como: cuerda de la circunferencia, segmentos circulares, arco de cuerda, faja central, razones trigonométricas, fórmulas de ángulo doble, etc. Este corte consiste en particionar la pizza a partir de dos cuerdas paralelas de la circunferencia: cuerda AB y cuerda EF . Estas cuerdas definen tres superficies: dos segmentos circulares, correspondientes al área comprendida entre las cuerdas y sus respectivos arcos de cuerdas (zona rosa y zona celeste en la figura 4) y una faja central, correspondiente al área comprendida entre ambas cuerdas (zona anaranjada en la figura 4). Para garantizar que las tres superficies tienen igual área, se debe determinar el ángulo central que contiene a los arcos de cuerdas de los segmentos circulares correspondientes (α y β).

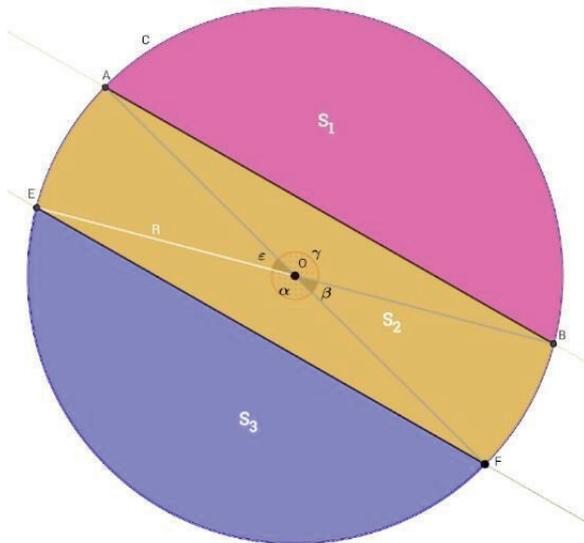


Figura 4. Segmentos y faja circulares (elaboración propia).

Por ejemplo, para calcular el área del segmento circular comprendido entre la cuerda AB y su arco (zona rosa en la figura 4 y denotada por S_1) se puede realizar la resta entre el área del sector circular delimitado por el ángulo γ (zona gris en la figura 5) y el triángulo OAB . El área del sector circular es $\frac{\pi R^2 \cdot \gamma}{360}$. Para

el área del triángulo OAB, es necesario recurrir a las razones trigonométricas en la circunferencia C. Así se obtiene que la altura h del triángulo $h = R \cos\left(\frac{\gamma}{2}\right)$ y la base $b = 2R \sin\left(\frac{\gamma}{2}\right)$. El área del triángulo OAB es entonces $\frac{R^2 \cdot \sin(\gamma)}{2}$. El área del segmento circular es $S_1 = R^2 \left(\frac{\pi \cdot \gamma}{360} - \frac{\sin(\gamma)}{2}\right)$. Para calcular el área de la faja central, esto es, el área comprendida entre la cuerda AB y la cuerda EF (denotada por S_2) se puede sumar el área del triángulo OAB, del triángulo OEF, del sector circular delimitado por el ángulo α y del sector circular delimitado por el ángulo β (figura 6). Ambos triángulos, al igual que los dos sectores circulares, tienen la misma área. Entonces, para obtener el área de la faja central, basta calcular el área de un triángulo y de un sector circular.

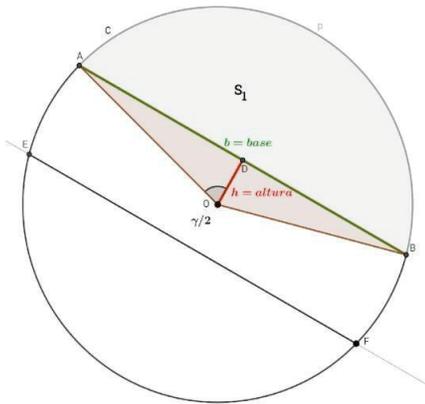


Figura 5. Cálculo del área del segmento circular (elaboración propia).

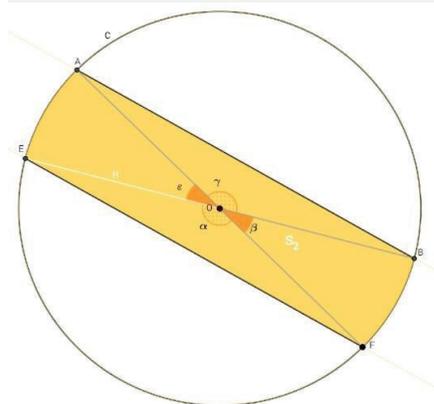


Figura 6. Cálculo del área de la faja central (elaboración propia).

El cálculo de las tres áreas conduce a un sistema de dos ecuaciones con dos incógnitas. Las incógnitas son los ángulos gamma y épsilon. La resolución de este sistema permite obtener algebraicamente, que gamma y épsilon son complementarios. Sin embargo, se necesita de un método de aproximación numérica para poder establecer la amplitud de los ángulos y poder calcular las áreas.

4.2.5. Divisiones de la pizza en forma de "T": segmento circular y mediatriz de la cuerda que lo contiene

Este corte requiere nociones tales como: segmento circular, cuerda, arco de cuerda y las propiedades entre estos elementos geométricos en relación al diámetro de la circunferencia. El corte consiste en particionar la pizza a partir de un segmento circular, correspondiente al área comprendida entre la cuerda AD y su respectivo arco de cuerda (superficie verde en la figura 7) y dos partes de los semicírculos (superficie rosa y celeste en la figura 7). Para garantizar que las tres superficies tienen igual área, se debe determinar el ángulo central AOC (β) que contiene al arco de cuerda AD, luego, el punto E de intersección entre el diámetro de la circunferencia C y la cuerda AD.

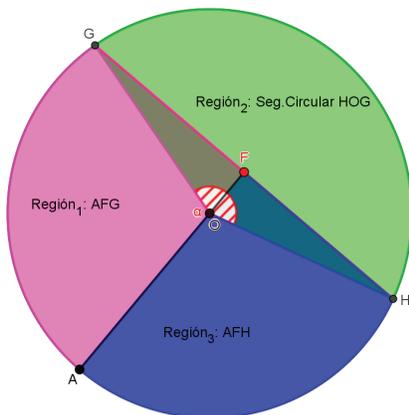


Figura 7. Corte en forma de "T" (elaboración propia).

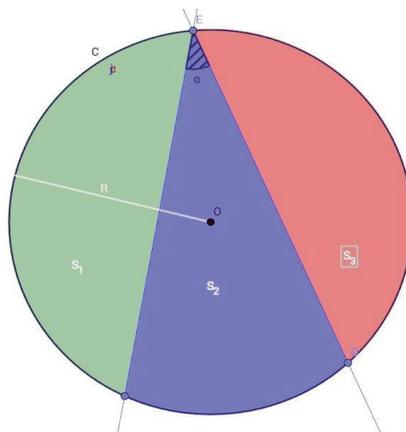


Figura 8. Corte en forma de "V" (elaboración propia)

4.2.6. Divisiones de la pizza en forma de "V": ángulo inscrito en la circunferencia

Este corte involucra nociones tales como: ángulo inscrito en la circunferencia, cuerda, ángulo central, sector circular, segmento circular, etc., y las propiedades entre estos elementos. El corte consiste en determinar la amplitud del ángulo alfa inscrito en la circunferencia (figura 8), de forma que las tres superficies

delimitadas por este ángulo tengan igual área. Las superficies son dos segmentos circulares (el verde y el rojo en la figura 8) formados por las cuerdas que forman el ángulo α y sus respectivos arcos de cuerda, y una superficie central (color azul en la figura 8) comprendida entre las cuerdas que conforman el ángulo. Para garantizar que las tres superficies tienen igual área, se debe determinar el ángulo α , lo que conduce a un profundo trabajo geométrico y algebraico.

4.2.7. Divisiones de la pizza a partir del Mosaico de disco monoédrico

Este corte involucra el concepto de teselado, que implica cubrir completamente una superficie plana con figuras idénticas, de forma tal que no se superpongan y luego, agruparlas de la forma conveniente de manera de obtener tres regiones de igual área. Haddley y Worsley (2015) proponen teselaciones del círculo a partir de diferentes figuras planas construidas con curvas (incluyendo rectas). Si bien, el trabajo de estos matemáticos no refiere exactamente al problema de la pizza, sí es posible utilizar la técnica del teselado para generar un nuevo corte (figura 9).

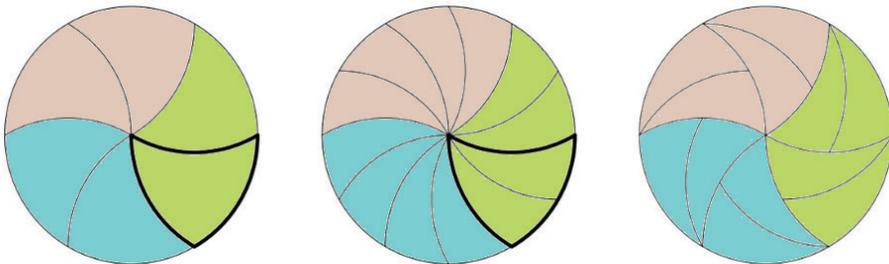


Figura 9. Mosaico de disco monoédrico (Haddley y Worsley, 2015)

Usando esta técnica, las nociones matemáticas y procedimientos puestos en juego se diferencian de los cortes anteriores. En este caso, se requiere cortar la pizza no sólo en superficies de igual área, sino también, con formas idénticas.

5. METODOLOGÍA DE LA INVESTIGACIÓN

Este trabajo reporta resultados de una investigación de tipo exploratoria. Se analizan las respuestas que aportaron ocho profesores al estudio del *problema de la pizza*, y las respuestas que generaron al mismo problema los siguientes recursos de IAG: ChatGPT 3.5 (ChatGPT), Google Bard y Bing Chat (en sus tres modalidades: más creativo, más equilibrado y más preciso). Analizamos diferencias y similitudes de estas resoluciones en términos de las siguientes categorías de análisis: el tipo de corte propuesto, el sistema de representación predominante en el cual se realiza la resolución, el nivel de validación matemática y el tipo de resolución (correcta, incorrecta o inconsistente). Los chatbots considerados basan la generación de texto en distintos modelos de lenguaje: GPT 3.5 (Brown *et al.*, 2020) en el caso de ChatGPT; GPT 4 (OpenAI, 2023) en el caso de Bing Chat y PaLM-2 (Chowdhery *et al.*, 2023) en el caso de Google Bard. Los modelos subyacentes a estos chatbots si bien se encuentran todos basados en una misma arquitectura de aprendizaje profundo, difieren principalmente en su escala y capacidades de generación de lenguaje, además de haber sido entrenados sobre subconjuntos de datos diferentes. La inclusión de múltiples chatbots tiene por objetivo analizar y comparar las respuestas generadas por cada uno de ellos, ya que la precisión, diversidad y capacidades en general de cada LLM varía. Se seleccionaron estos tres chatbots de propósito general por ser los que se encuentran disponibles de manera libre y ser los más ampliamente adoptados entre estudiantes y docentes.

El grupo de profesores se encontraba realizando un curso de formación continua que correspondía a la resolución de problemas de matemática utilizando recursos tecnológicos como GeoGebra, planilla de cálculo, programación por bloques y recursos provenientes de IAG. El curso basa su metodología de enseñanza en ABP pues, entre otras características, propone a los profesores resolver los problemas directamente a través de los recursos tecnológicos. Para cada recurso se propusieron actividades individuales y grupales. Los profesores de matemáticas, tenían edades entre 25 y 50 años. Todos se desempeñaban en los diferentes años del nivel secundario de Argentina y en diversas ciudades del país. La actividad final del curso consistió en la resolución, de manera individual, de un problema de matemática usando todos los recursos estudiados y con un análisis de las ventajas y desventajas de cada uno de ellos. A cada estudiante se le asignó un problema diferente y dispusieron de diez días para presentar el trabajo, de forma escrita y oral. Ninguno de los profesores conocía previamente

el problema. Además de resolverlo, debían diseñar una propuesta de enseñanza usando el problema asignado, detallando el nivel escolar destinatario de la propuesta, las nociones matemáticas que pretendían estudiar, las modificaciones realizadas al mismo y cómo gestionaría la clase. Es requisito que esa propuesta contemple al menos dos recursos tecnológicos a usar con los estudiantes.

Parte de las decisiones metodológicas fue no aportar, ni a los profesores ni a los chatbots especificaciones sobre el formato de la pizza (circular o no). Justamente, se buscó remarcar en este punto que la IA no tiene conocimientos sobre supuestos comunes o asunciones que normalmente hacemos las personas sobre las características de los objetos.

A los profesores se los identifica como P_i , donde i indica el número de cada uno. Respecto a las soluciones aportadas por los chatbots, las respuestas de estos recursos de IAG fueron resultantes de varias interacciones con los mismos. La formulación del *problema de la pizza* fue utilizado como *prompt* inicial para los chatbots considerados en el estudio. La IAG se caracteriza por generar un texto nuevo cada vez que se le solicita, por lo tanto las respuestas de los chatbots varían para la misma pregunta, siendo por naturaleza no determinísticos. Con el fin de contemplar esta aleatoriedad en las respuestas se analizaron tres respuestas para cada *prompt* utilizado en un chatbot dado. Las respuestas al *prompt* inicial, en general, indicaron que el corte en forma de sector circular no es el único posible y ofrecieron una serie de alternativas. A partir de una respuesta inicial, se continuó preguntando sobre la validación matemática de dichos métodos de corte, utilizando el siguiente *prompt*: *¿Podrías demostrar matemáticamente que las áreas son equivalentes si se corta la pizza usando estos métodos?*

Con el objeto de obtener una representación gráfica de los cortes que pueden realizarse con cada método alternativo, se pidió a los chatbots que generen un código fuente (sin especificar el lenguaje) que permita graficarlos, utilizando el siguiente *prompt*: *¿Podrías generar el código necesario para graficar los cortes de pizza con estos métodos?* Las representaciones gráficas obtenidas son el resultado de ejecutar los códigos ofrecidos como respuesta en cada caso que, aunque no fue indicado en el *prompt*, en todos los casos fue en lenguaje Python. Las respuestas, tanto de los profesores como de los chatbots, se analizaron mediante la tabla 1, que permitió generar de manera inductiva (Isaza, 2002) las categorías y sus correspondientes subcategorías de análisis.

Tabla 1. Análisis de las respuestas

Profesor (P _i)/ IAG	C ₁ : Tipo de corte propuesto	C ₂ : Marco preponderante	C ₃ : Nivel de validación	C ₄ : Tipo de resolución	Representación gráfica del corte propuesto
---------------------------------	--	--------------------------------------	--------------------------------------	-------------------------------------	--

En la primera columna se identificó el generador de la respuesta, que puede ser un profesor (P_i) o un recurso de IAG (ChatGPT 3.5, Bard, Bing Chat). En la segunda columna se colocó la categoría Tipo de corte propuesto (C₁). Sus correspondientes subcategorías son las siguientes:

- **C_{1.1}: Sector circular (y sus variantes)**: Consiste en partir la pizza mediante tres sectores circulares cuyo ángulo central es de 120° (figura 1). Sus variantes refieren a superficies circulares con ángulo central de 120° pero que sus radios no son necesariamente rectos.
- **C_{1.2}: Coronas circulares**: Considera dos coronas circulares y un círculo central (los tres concéntricos) (figura 2).
- **C_{1.3}: Rectas paralelas**: Consiste en particionar la pizza a partir de dos cuerdas paralelas de la circunferencia, definiendo mediante ellas tres superficies: dos segmentos circulares, correspondientes al área comprendida entre las cuerdas y sus respectivos arcos de cuerdas y una faja central, correspondiente al área comprendida entre ambas cuerdas (figura 4).
- **C_{1.4}: Corte "T"**: Consiste en particionar la pizza a partir de un segmento circular, correspondiente al área comprendida entre una cuerda y su respectivo arco de cuerda; y luego, dos partes de los semicírculos (figura 7).
- **C_{1.5}: Ángulo inscrito en la circunferencia**: Consiste en determinar la amplitud de un ángulo alfa inscrito en la circunferencia, de forma que las tres superficies delimitadas por este ángulo tengan igual área (figura 8).
- **C_{1.4}: Por peso**: Refiere a cortar una pizza real y tratar de lograr las equivalencias de pesos, más que de áreas. Si bien este corte fue propuesto por uno de los profesores, no se ajustaría al requerimiento de figuras de igual área ya que, dos o más figuras pueden tener igual área pero no el mismo peso.
- **C_{1.5}: Una corona circular y dos semicírculos**: Corresponde a considerar una corona circular y un círculo central (ambos concéntricos) particionado en dos mitades.

- **C_{1.6}: Espiral:** Supone un corte que parte desde el centro de la pizza y se corta en espiral hacia afuera en tres secciones iguales.
- **C_{1.7}: Rectángulos:** Supone una pizza rectangular y consiste en dividirla en tres rectángulos semejantes de igual área.
- **C_{1.8}: Arbitrario:** Propone particionar la pizza en representaciones figurativas, como por ejemplo, hacer un corte en forma de estrella, con cinco puntas; hacer un corte en forma de flor, con seis pétalos; hacer un corte en forma de animal, como un perro.
- **C_{1.9}: Mosaico de disco monoédrico:** Involucra el concepto de teselado. Esta noción implica cubrir completamente una superficie plana con figuras idénticas, de forma tal que no se superpongan (figura 9).
- **C_{1.10}: Teorema de la pizza:** Utiliza el teorema de la pizza comentado en el apartado 2.1, para cortar la pizza en sectores y determinar la igualdad de la suma de sus áreas.

En la tercera columna de la tabla 1 se hizo referencia a la categoría Sistema de representación preponderante (C_2) en el que se realizó la solución. Estos sistemas de representación se definieron en relación a las técnicas que se proponen para poder desarrollar el corte seleccionado y justificar así la respuesta propuesta. Estas técnicas pueden incluir más de un sistema, y las subcategorías definidas son las siguientes:

- **C_{2.1}: Algebraico:** Desarrollo de fórmulas para mostrar las áreas de cada porción de pizza en forma general, y para mostrar la igualdad de área entre ellas.
- **C_{2.2}: Geométrico:** Construcciones de la geometría plana y establece la validación a partir del análisis de propiedades geométricas.
- **C_{2.3}: Aritmético:** Ejemplos particulares para calcular áreas específicas y probar la igualdad entre ellas numéricamente.
- **C_{2.4}: Coloquial:** Explicaciones en palabras tomadas del lenguaje común. Por ejemplo: "Si cortas una pizza en seis partes iguales, cada porción representa un sexto de la pizza. Si tomas dos de estas porciones, tendrás $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$ de la pizza. Por lo tanto, tres personas pueden tener dos porciones cada una, lo que equivale a un tercio de la pizza para cada una".
- **C_{2.5}: Gráfico:** Propuesta de esquemas a mano alzada y a partir de ellos se realiza la validación.

- **C_{2.6}: Empírico:** Corte de una pizza real, y medir de alguna manera (peso, o con regla) la igualdad de los trozos.

La cuarta columna de la tabla 1 se refiere a la categoría Nivel de validación (C₃). Se analizó el procedimiento propuesto para validar matemáticamente que las tres superficies de la pizza generadas por el corte seleccionado tienen igual área. En el caso de las respuestas obtenidas (tanto de los profesores, como de las IAG), encontramos validaciones que recurren no sólo a este proceso de razonamiento, sino a otras técnicas como por ejemplo a la propuesta de casos particulares y resoluciones aritméticas para justificar las mismas. Las subcategorías generadas son las siguientes:

- **C_{3.1}: Nula:** La respuesta no es acompañada por ninguna justificación.
- **C_{3.2}: Baja:** Se realizan algunos intentos de validación pero se reducen a ejemplos numéricos.
- **C_{3.3}: Media:** Se advierte un intento de justificación matemática centrada en calcular algebraicamente el área de cada superficie (o región) de la pizza, y la demostración de igualdad de las áreas de cada superficie; pero no logra una generalización, y la propuesta queda vinculada al caso concreto de estudio.
- **C_{3.4}: Alta:** Se realiza un proceso de justificación matemático centrado en calcular algebraicamente el área de cada superficie (o región) de la pizza, así como la demostración de igualdad de las áreas de cada superficie, independientemente del tamaño de la pizza (o equivalentemente, de la longitud del radio de la circunferencia).

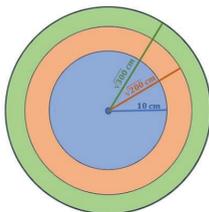
En la quinta columna de la tabla 1 se registró el Tipo de resolución (C₄), esto hace referencia a los razonamientos explicitados en el desarrollo de la resolución y al tipo de resultados que estos permiten obtener. No se trató de determinar si la respuesta es completa o incompleta; sino si el procedimiento empleado permitió obtener o no (aunque sea parcial) resultados correctos, incorrectos o si se advirtieron inconsistencias, independientemente del tipo de marco empleado y del nivel de validación ofrecida. Por consiguiente, las subcategorías generadas fueron las siguientes:

- **C_{4.1}: Incorrecto:** Se confundieron nociones matemáticas, se obtuvieron resultados incorrectos en los cálculos, se asumió la equivalencia de

- identidades que no lo son, entre otros. Por ejemplo, confundir perímetro de la circunferencia con área del círculo, asegurar que $\frac{2}{3} + 2 = 1$; igualar expresiones que no son ciertas por ejemplo: $\frac{3}{4} \pi = \frac{1}{3} \pi$; entre otros.
- **C_{4.2}: Correcta:** No se cometió ningún error de procedimientos incluso cuando la respuesta está incompleta.
 - **C_{4.3}: Inconsistente:** Deriva de en una respuesta incorrecta, pero no es consecuencia de un error de cálculo o procedimiento, sino que se vincula a la idea de contradicción matemática. Se distinguió este tipo de error por su posible relación a la forma de construcción de las respuestas de la IAG, carente de razonamiento deductivo. Por ejemplo, falta de consistencia cuando se propone un tipo de corte, mientras que el procedimiento y explicaciones refieren a otra forma de corte; incluso, en el caso de los chatbot, la gráfica que se genera al ejecutar el código Python propuesto presenta una construcción diferente a los tipos de partición antes referidos.

En la sexta columna de la tabla 1 se indicó un esquema, figura o gráfico relativo al corte propuesto sea por los profesores o por los chatbots. En el caso de los chatbots, los gráficos se obtuvieron a partir de ejecutar el código fuente aportado. A continuación, se coloca un fragmento de respuesta, correspondiente al Profesor P1.

El corte tradicional en forma de sector circular es el más extendido y fácil pero no es el único. Por ejemplo, otra manera de poder dividir la pizza es tres partes iguales es:

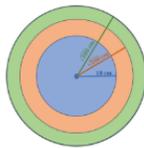


Las tres porciones, azul, naranja y verde tienen la misma área por ende son porciones iguales. La porción azul tiene un radio de 10 cm , su área es: $\pi \cdot 10^2 = 100\pi$. La porción naranja tiene un radio de $\sqrt{200}\text{ cm}$, su área es: $\pi \cdot \sqrt{200}^2 = 200\pi$ menos la porción azul es $200\pi - 100\pi = 100\pi$. La porción verde tiene un radio de $\sqrt{300}\text{ cm}$, su área es: $\pi \cdot \sqrt{300}^2 = 300\pi$ menos la porción naranja es $300\pi - 200\pi = 100\pi$

Figura 10. Fragmento de solución aportada por P1.

P1 propuso el corte a partir de coronas circulares y círculo central (los tres concéntricos): supuso un caso particular, considerando un radio específico para cada corona circular y para el círculo central. Considerando estos valores de radios, calculó el área para cada una de las tres superficies: las dos coronas (verde y naranja) y para el círculo (azul). Logró concluir que las tres figuras tienen igual área. La fila de la tabla 1 correspondiente a P1 se completa entonces de la siguiente manera (tabla 2):

Tabla 2. Ejemplo de análisis de respuesta de P₁

Profesor (P ₁)/ IAG	C ₁ : Tipo de corte propuesto	C ₂ : Sistema de representación preponderante	C ₃ : Nivel de validación	C ₄ : Tipo de resolución	Representación gráfica del corte propuesto
P ₁	Sector circular	Aritmético	Validación baja: se sostiene sobre un ejemplo numérico.	Correcta	

También se construyó una tabla para cada una de las categorías. En la primera columna se colocó la categoría considerada (sea C_1 : Tipo de corte propuesto; C_2 : Sistema de representación preponderante; C_3 : Nivel de validación o C_4 : Tipo de resolución, según corresponda). De la segunda a la séptima columna, respectivamente, se consideró: ChatGPT 3.5; Bard; Bing Chat Más Creativo; Bing Chat Más Equilibrado; Bing Chat Más Preciso; Profesores. En relación a los Profesores, es importante destacar la diferencia respecto a la primera tabla: en la tabla 1 se colocó una fila para cada uno de los ocho profesores, mientras que en las tablas para cada categoría, se contabilizó por cantidad de respuestas del total de los profesores (sin distinguir entre cada uno de los ocho). Por ejemplo, para la categoría C_1 , en la columna relativa a *Profesores* se contabilizó el total de respuestas identificadas para cada uno de los cortes (es decir, las subcategorías). Estas tablas específicas a las categorías, se computaron haciendo uso de las frecuencias relativas por columnas. Se presenta como ejemplo de este procesamiento de datos, la tabla y las frecuencias relativas correspondientes a la categoría C_1 (tabla 3).

Tabla 3. Análisis con frecuencias relativas, categoría C_1

C_1 : Tipo de corte propuesto	ChatGPT 3.5	Bard	Bing Chat			Profesores
			Más Creativo	Más Equilibrado	Más Preciso	
Sector circular (y sus variantes)	0.4	0.6	1	0	0.5	0.55
Coronas circulares	0	0	0	0	0	0.15
Rectas paralelas	0	0	0	0	0	0.1
Corte "T"	0	0	0	0	0	0.05
Ángulo inscripto en circunferencia	0	0	0	0	0	0.05
Por peso	0	0	0	0	0	0.05
Corona circular y dos semicírculos	0	0	0	0	0	0.05
Espiral	0.2	0	0	0	0	0
Rectángulos	0.4	0	0	0	0	0
Arbitrario	0	0.2	0	0	0	0
Mosaico de disco monoédrico	0	0.2	0	0	0.5	0
Teorema de la pizza	0		0	1	0	0

6. RESULTADOS Y DISCUSIÓN

Las respuestas ofrecidas por los profesores y los chatbots fue de 36: 20 realizadas por los profesores y 16 por los chatbots. El análisis de cada una de ellas y los datos obtenidos para las diferentes categorías de la tabla 1 se detallan a continuación.

C₁: Categoría 1. Tipo de corte propuesto: Se identificaron doce tipos de cortes diferentes, siendo el de mayor frecuencia el denominado *sector circular (y sus variantes)*. Este corte es el que habitualmente se elige para dividir una pizza en tres partes, aun cuando no sea tan estricto en demostrar que son figuras de igual área. Se destaca la variedad de tipo de cortes propuestos por los profesores (7 tipos de cortes diferentes), en comparación con los propuestos por los chatbots, quienes realizaron propuestas de a lo sumo tres tipos de cortes diferentes, pues *corte en espiral* y *teorema de la pizza*, se reducen en realidad al corte circular. ChatGPT propuso un corte en rectángulos, para una pizza a la que supuso rectangular, y lo validó correctamente en el sistema algebraico. Sorprende que es el único que lo hizo (de chatbots y profesores), siendo que el enunciado no propone una pizza circular específicamente. En las respuestas de los profesores es de esperar que supongan una pizza circular, debido a que se sostiene sobre una idea adquirida socialmente cada vez que se compra o se cocina una pizza. Por el contrario, los chatbots no adquieren explícitamente supuestos comunes, sino que responden en base a un contexto (las palabras dentro de la conversación) y las relaciones entre palabras aprendidas en base a los textos observados. Si bien hay alta probabilidad de que en estos textos se asocie la pizza a formatos circulares, de ahí que la mayor parte de las respuestas también lo hace, la solución rectangular no queda excluida de las posibles respuestas. Este ejemplo puso en evidencia la importancia de dar contexto a una pregunta para un chatbot, en este caso especificar la forma de la pizza, ya que eso orienta la predicción del texto resultante. Lo mismo sucede en el caso que sugirió cortes de forma arbitraria (como en forma de estrella, flor de seis pétalos y hasta dinosaurios), sin advertir el fin matemático del enunciado, que se acota con el segundo prompt donde se requiere específicamente una demostración (en cuyo caso descarta que la misma sea alcanzable para cortes arbitrarios). En la figura 11 se presenta un gráfico de barras sintetizando los tipos de corte propuesto según los chatbots y los profesores.

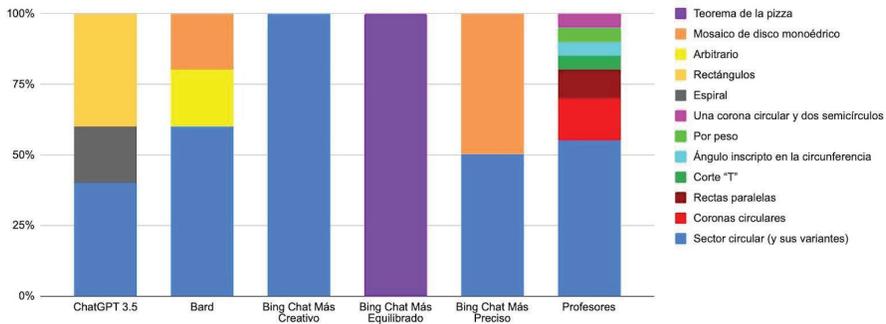


Figura 11. Cortes propuestos por profesores y chatbots.

C₂: Categoría 2. Sistema de representación preponderante: Se observó mayor variedad de sistemas de resolución en las propuestas de los profesores que en las de los chatbots. El sistema de resolución más frecuente propuesto por los chatbots fue el coloquial, mientras que entre los profesores se advirtió una misma proporción entre los tipos de resolución gráfico, geométrico y numérico (figura 12). El nivel algebraico, que matemáticamente sería el más indicado para resolver este tipo de problemas, fue propuesto por un único chatbot (ChatGPT) para algunas de sus resoluciones, y por una muy baja proporción de los profesores. En el caso de estos últimos, preocupa que el sistema de representación preponderante para su resolución no haya sido el algebraico, debido a que son quienes forman a nuestros estudiantes de secundaria en matemática. En el caso de los chatbots, se esperaba también un mayor uso del sistema algebraico, pues es el sistema de representación que tiene mayor potencial para resolverlo. Sin embargo, que hayan usado el sistema de representación coloquial se condice con su creación. Este resultado brinda un espacio de discusión en los colegios, con los estudiantes para evitar considerar a los chatbot perfectos y omniscientes (que lo saben todo).

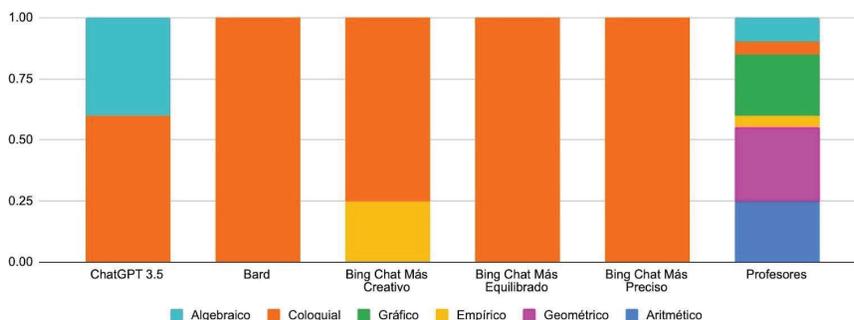


Figura 12. Sistema de representación utilizados por profesores y chatbots.

C₃: Categoría 3. Nivel de validación: El siguiente gráfico (figura 13) muestra la distribución de los tipos de validaciones matemáticas. Se destaca la ausencia de validación en las propuestas de la mayoría de las respuestas obtenidas de los chatbots. ChatGPT es el único que proporcionó respuestas donde el tipo de validación fue calificada como *alta*. En relación a las respuestas proporcionadas por los profesores, los resultados fueron más variados, aunque el tipo de validación más típico es *nulo* o *bajo*. En el caso de los chatbots, fue necesario continuar preguntando sobre la validación matemática de dichos métodos de corte, pues de lo contrario no las realizaban. En cambio, los profesores de matemática en servicio, y a pesar de no haberles solicitado explícitamente validar la equivalencia de áreas, la mayoría de ellos ofreció al menos una validación, aunque la preponderancia fue de tipo baja. Esto se debió, de forma análoga a la categoría anterior, a que los chatbots no adquieren explícitamente supuestos comunes. En cambio, esto sí ocurre entre las personas, y en particular, entre los profesores de matemática: dentro de la formación de un profesor de Matemática, está instalado el supuesto que toda afirmación debe demostrarse.

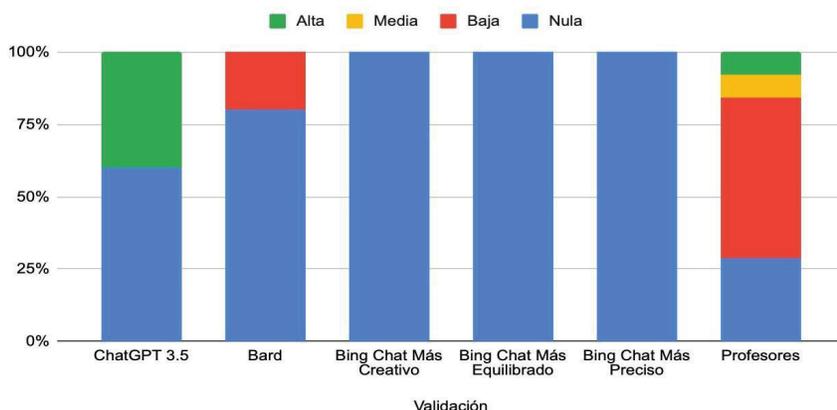


Figura 13. Validaciones matemáticas de profesores y chatbots.

C₄: Categoría 4. Tipo de Resolución. En el caso de los profesores las respuestas proporcionadas fueron siempre correctas: ninguno de ellos cometió errores o incurrió en inconsistencias. En algunos casos estaban incompletas, pero el procedimiento realizado era correcto. Contrariamente, para las respuestas proporcionadas por los chatbots se destaca el alto nivel de respuestas incorrectas o inconsistentes; solo ChatGPT proporcionó alguna respuesta correcta. Para la subcategoría *inconsistente*, se identificaron respuestas como la presentada en la figura 14, que fue proporcionada por el ChatGPT. En esta respuesta no hay consistencia entre el nombre del corte que propone el ChatGPT, en *espiral*, con la forma en que se propone resolver, asumiendo *sectores circulares*, y la representación gráfica resultante al ejecutar el código Python: *coronas circulares*.

Supongamos que cortamos la pizza en tres partes iguales mediante una espiral que parte desde el centro y llega al borde exterior.

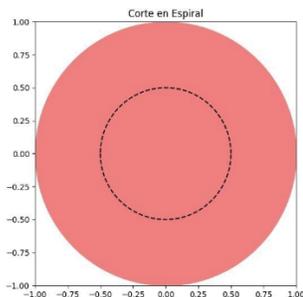


Figura 14. Respuesta inconsistente generada por el ChatGPT 3.5

La figura 15 sintetiza la distribución del tipo de respuesta (correcta, incorrecta e inconsistente) ofrecidas por chatbots y profesores.

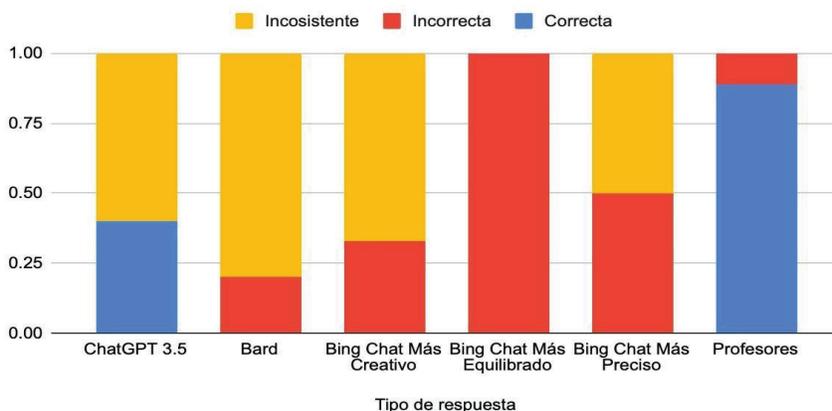


Figura 15. Resoluciones realizadas por profesores y chatbots.

En los trabajos descritos en la sección de antecedentes, se alienta el uso de recursos provenientes de la IAG como asistentes y herramientas para los educadores y los estudiantes, siempre y cuando se analicen críticamente las respuestas que ofrecen. En este trabajo, se profundizó sobre la variedad de resoluciones, sistemas de representación preponderante, nivel de validación matemática y tipo de resolución. Estos resultados reafirman las conclusiones obtenidas previamente en

Parra *et al.* (2024), donde se identificaron y categorizaron los errores en las respuestas de los mismos tres chatbots a un problema de geometría. Estos trabajos permitieron obtener una mayor comprensión de las capacidades de los chatbots en contraposición con los profesores de matemática.

7. CONCLUSIONES

En este trabajo presentamos resultados de explorar, analizar y comparar las respuestas que ofrecen los modelos de IAG con las respuestas que propusieron profesores de matemática ante un problema geométrico. Para cada respuesta analizamos diferencias y similitudes de las resoluciones a partir de las siguientes categorías: el tipo de corte propuesto, el sistema de representación predominante, el nivel de validación matemática y el tipo de resolución (correcta, incorrecta o inconsistente).

Respecto del tipo de corte, se observó una menor variedad en las soluciones ofrecidas por los chatbots en comparación a los profesores. Estas soluciones responden, además, al tipo de corte más usual, que es el de sectores circulares y sus variantes. Si bien eventualmente un chatbot podría responder con alternativas poco comunes (como ocurrió con la solución que asumió una pizza rectangular), las soluciones clásicas son las más probables por ser las que mayor presencia tienen en los textos. Los modelos de IAG están guiados por los datos que se usaron para su entrenamiento y la distribución de los mismos, donde hay un desbalance que favorece a las soluciones más difundidas o extensamente discutidas en la literatura. La falta de diversidad en las respuestas conlleva un riesgo asociado a los LLMs, que es el de favorecer y perpetuar teorías dominantes por sobre otras minoritarias o emergentes (Bender *et al.*, 2021).

Un caso interesante para el análisis se dio con la respuesta de ChatGPT que mencionó una pizza rectangular. Si bien, en el enunciado del problema no se explicita que la pizza adopta la forma de un círculo, la inclusión de la noción “sector circular” condujo a todos los profesores a asumir un círculo. En cambio, en el caso de los chatbots la mención a “sector circular” contextualiza el problema pero no restringe la solución a una forma de pizza en particular, ya que la respuesta puede construirse en base a conceptos de menor probabilidad de aparición.

La presencia de inconsistencias entre la alternativa de corte propuesto, su validación matemática y la representación gráfica, se observó como un elemento distintivo de las respuestas de los chatbots, que naturalmente no aparece en

las respuestas de los profesores. Estas inconsistencias pueden atribuirse al mecanismo de generación de textos de la IAG, por un lado, pero más esencialmente a la brecha entre la tarea para la que están entrenados los LLMs (predecir palabras) y la tarea para la cual se aplican (razonamiento matemático en este caso). En el caso de los profesores, no se identificaron inconsistencias aunque sí respuestas incorrectas e incluso, incompletas. Esto podría deberse posiblemente a la formación e incluso, la experiencia de estos profesores, pues se trata de profesores de matemática en servicio, que se desempeñan, la mayoría de ellos, en el nivel secundario argentino donde las tareas escolares habitualmente admiten una única respuesta y centradas en pocos saberes matemáticos.

Los LLMs están contruidos para predecir la siguiente palabra en una secuencia de palabras previas, de manera que el texto se construye prediciendo y agregando la palabra más probable de acuerdo a las anteriores. Las inconsistencia surgen por múltiples razones. El mecanismo de predicción de la próxima palabra es inherentemente hacia adelante, no hay un proceso de revisión, una predicción errada condiciona la construcción del texto posterior por lo que puede desviarse en cualquier punto hacia una solución incorrecta. La predicción está sujeta a un contexto que no necesariamente está descrito de manera precisa y completa, ya que se construye a medida que se incorporan nuevas palabras, que a su vez tienen solo una cierta probabilidad asociada. La falta de precisión en la descripción del contexto lleva a predicciones que se alejan del problema planteado y del texto parcialmente construido, derivando en dichas inconsistencias.

La diferencia con el razonamiento matemático que pueden desarrollar en este caso los profesores radica en que la generación de texto de una IAG no está dirigida por un objetivo que le posibilite el planear por adelantado para alcanzarlo, ni tiene un “panorama completo del problema” incluido tal objetivo, una completa comprensión de los elementos matemáticos relevantes e incluso supuestos comunes. En un análisis de las capacidades en razonamiento matemático de GPT-4 (modelado usado por Bing Chat), Bubeck *et al.* (2023) mostraron a través de experimentos que el modelo tiene un alto nivel para elegir el argumento o camino correcto hacia la solución y es capaz de desarrollar procedimientos matemáticos, aunque frecuentemente comete errores (incluyendo errores aritméticos) y presenta deficiencias significativas en términos de razonamiento crítico. La carencia en la capacidad de examinar críticamente cada paso del argumento viene asociada a la imposibilidad de volver atrás cuando se descubre que un paso es incorrecto y modificar el argumento en consecuencia. Por estas

marcadas diferencias, que son intrínsecas a los mecanismos de generación de la IAG, McCoy *et al.* (2023) destacan la importancia de no ver a los LLMs como resolutores de problemas matemáticos, sino más bien como un sistema estadístico de predicción de palabras que se utiliza para resolver problemas matemáticos. Entonces, los fracasos pueden entenderse directamente en términos de un conflicto entre la tarea de predicción de palabras y la de resolución de problemas matemáticos.

A lo anterior se suma que, incluso los LLMs multimodales más avanzados muestran limitaciones al abordar problemas geométricos, produciendo varias veces descripciones inexactas de figuras geométricas debido a los desafíos para comprender con precisión las relaciones entre elementos fundamentales en figuras geométricas como puntos y líneas e interpretar correctamente nociones simples como el grado de un ángulo. Gao *et al.* (2023) argumentan que la razón subyacente de esto puede ser el hecho de que los LLM multimodales generalmente se entrenan con imágenes y descripciones de dominio general, y la capacidad de comprender dicha semántica difiere significativamente de la requerida para el razonamiento geométrico. Estos resultados se condicen con los obtenidos en Parra *et al.* (2024) respecto a la categorización de errores cometidos por estos LLMs al resolver un problema geométrico perteneciente a las Olimpiadas Matemáticas Argentinas. Se identifican, para el caso del *problema de la pizza*, diversos errores vinculados a la confección de construcciones y a inconsistencias entre deducciones.

Al comparar las respuestas aportadas al problema de la pizza por los modelos de IAG empleados para esta investigación, con las respuestas aportadas por los profesores, destacamos que las de los modelos de IAG, en su mayoría, son inconsistentes, poco variadas y la validación es baja en términos de rigurosidad matemática. Los resultados obtenidos indican que no resulta apropiado emplear estos modelos como recurso en la resolución de problemas de geometría, sin una reflexión crítica tanto por parte de los profesores como de los estudiantes. Un posible uso en el aula de estos chatbots podría ser el análisis de las respuestas obtenidas, lo que permitiría potenciar el aprendizaje de la geometría (propiedades, características, construcciones en el plano, etc.) de forma crítica, por ejemplo, distinguiendo cuando es posible (o no) aplicar un teorema (lema, corolario, etc.).

Finalmente, respecto al análisis del problema de la pizza y al uso de los chatbots como recursos posibles de nutrir la práctica docente, concluimos que el problema tiene múltiples potencialidades para el estudio de la geometría en la escuela secundaria usando recursos de la IAG de forma crítica.

REFERENCIAS

- Bender, E. M., Gebru, T., McMillan-Major, A., y Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT'21)*, 610-623, Virtual Event, Canada.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ..., y Amodei, D. (2020). Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20)*, Vancouver, Canada.
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., y Zhang, Y. (2023). *Sparks of artificial general intelligence: Early experiments with GPT-4*. CoRR abs/2303.12712.
- Burgos Cantos, M., y Cao, E. (2023). Herramientas tecnológicas más usadas en la enseñanza de la matemática. *Suplemento CICA Multidisciplinario*, 7(16), 166-181. <https://doi.org/10.60100/scicam.v7i016.117>
- Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., ..., y Fiedel, N. (2023). PaLM: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240), 1-113.
- Camacho Machín, M. (2015). El desarrollo de la actividad matemática con estudiantes de bachillerato mediante el uso de la tecnología para la resolución de problemas: algunos ejemplos. En: M. V. Becerra Sepúlveda (Ed.), *Construcción de modelos matemáticos y resolución de problemas* (pp.43-67). Madrid, Ministerio de Educación y Ciencia.
- Corica, A. Parra, V., Sureda, P., Schiaffino, S., y Godoy, D., (2024). Fractal de Koch: análisis de respuestas de IA generativa y un profesor de matemática, *TE&ET*, 37, e8.
- Dirección General de Escuelas de la Provincia de Buenos Aires. (2006). *Diseño Curricular para la Educación Secundaria Matemática*. Gobierno de la Provincia de Buenos Aires.
- Espinoza-Benavidez, N. E., y Triminio-Zavala, C. M. (2018). Implementación de la metodología Aprendizaje Basado en Problemas (ABP), en la unidad de Geometría Plana. *Revista Científica Estelí*, 26, 65-76. <https://doi.org/10.5377/farem.v0i26.6430>
- Espinoza Melo, C., y Sánchez Soto, I. (2014). Aprendizaje Basado en Problemas para enseñar y aprender estadística y probabilidad. *PARADIGMA*, XXXV(1), 103 - 128.
- Gallent-Torres, C., Zapata González, A., y Ortego Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. *RELIEVE*, 29(2). <https://doi.org/10.30827/relieve.v29i2.29134>

- Gao, J., Pi, R., Zhang, J., Ye, J., Zhong, W., Wang, Y., Hong, L., Han, J., Xu, H., Li, Z., y Kong, L. (2023). G-LLaVA: Solving geometric problem with multi-modal large language model. *arXiv*, vol. abs/2312.11370
- García, I. (2016). Uso de la metodología ABP (Aprendizaje Basado en Problemas) para desarrollar los contenidos del bloque de geometría en matemáticas. *Publicaciones Didácticas (E-Journal)*, 76, 443-449.
- Grisales-Aguirre, A. (2018). Uso de recursos TIC en la enseñanza de las matemáticas: retos y perspectivas. *Entramado*, 14(2), 198-214. <https://doi.org/10.18041/1900-3803/entramado.24751>
- Haddley, J. A., y Worsley, S. (2015). Infinite families of monohedral disk tilings. *ArXiv*. <https://doi.org/10.48550/arXiv.1512.03794>
- Isaza G. (2002). *Análisis, Interpretación y Construcción Teórica en la Investigación Cualitativa*. Centro de educación a distancia. Universidad de Manizales.
- Itzcovich, H. (2005). *Iniciación al estudio didáctico de la Geometría: de las construcciones a las demostraciones* (Vol. 3). Libros del zorzal.
- Kravitz, C., Stanley Ogilvy, S., Tan, K., Mozzochi, C. J., Upton, L. J., Rabinowitz, S., y Baum, J. D. (1967). Problems and Solutions. *Mathematics Magazine*, 40(3), 162-172. <https://doi.org/10.2307/2688484>
- Lozada Lozada, R. F., Lopez Aguayo, E. M., Espinoza Suquilanda, M. de J., Arias Pico, N. de J., y Quille Vélez, G. E. (2023). Los Riesgos de la Inteligencia Artificial en la Educación. *Ciencia Latina Revista Científica Multidisciplinar*, 7(5), 7219-7234. https://doi.org/10.37811/cl_rcm.v7i5.8301
- Molina-Patiño, E. K., Hernández-Martínez, M. A., Mayanquer-Barreto, J. C., Montenegro Reinoso, M. I., y Mora Grijalva, M. M. (2023). La metodología de aprendizaje activo “aprendizaje basado en problemas” en la formación de futuros maestros de matemáticas. *Ciencia Latina Revista Científica Multidisciplinar*, 7(3), 1276-1292. https://doi.org/10.37811/cl_rcm.v7i3.6271
- McCoy, R. T., Yao, S., Friedman, D., Hardy, M., y Griffiths, T. L. (2023). *Embers of autoregression: Understanding large language models through the problem they are trained to solve*. CoRR abs/2309.13638.
- Nelsen, R. B. (1993). *The proof without words: exercises in visual thinking*. The Mathematical Association of America.
- OpenAI (2023). GPT-4 technical report. *arXiv*, <https://doi.org/10.48550/arXiv.2303.08774> .
- Padilla-Doria, L., y Flórez-Nisperuza, E. (2022). El aprendizaje basado en problemas (ABP) en la educación matemática en Colombia. Avances de una revisión documental. *Redipe*, 11(2), 318-2. <https://doi.org/10.36260/rbr.v11i2.1686>

- Peralta Roncal, L. E., Gaona Portal, M. del P., Luna Acuña, M. L., y Bazán Linares, M. V. (2023). Las tecnologías de la información y la comunicación (TIC) en educación secundaria: Una revisión sistemática. *Revista Andina de Educación*, 7(1), 000711. <https://doi.org/10.32719/26312816.2023.7.1.1>
- Parra, V., Sureda, P., Corica, A., Schiaffino, S., y Godoy, D. (2024). Can generative AI solve Geometry problems? Strengths and weaknesses of LLMs for geometric reasoning in Spanish. *IJIMAI*, 8(5), 65-74. <https://dx.doi.org/10.9781/ijimai.2024.02.009>
- Polya, G. (1981). *Mathematical Discovery: On understanding, learning and teaching problem solving*. New York: Wiley.
- Rojas, C., y Sierra, T. (2020). Los problemas espaciales: una propuesta alternativa para enseñar geometría en la Educación Secundaria Obligatoria. *Educação Matemática Pesquisa*, 22(4), 593–602.
- Romo-Pérez, V., García-Soidán, J. L., Özdemir, A. S., y Leirós-Rodríguez, R. (2023). ChatGPT ha llegado ¿Y ahora qué hacemos? La creatividad, nuestro último refugio. *Revista de Investigación en Educación*, 21(3), 320-334. <https://doi.org/10.35869/reined.v21i3.4973>
- Sanabria-Navarro, J., Silveira-Pérez, Y., Pérez-Bravo, D., y de-Jesús-Cortina-Núñez, M. (2023). Incidences of artificial intelligence in contemporary education. *Comunicar*, 77, 97-107. <https://doi.org/10.3916/C77-2023-08>
- Schoenfeld, A. (1992). Learning to think mathematically: problem solving, metacognition and sense making in Mathematics. In D. Grouws (Ed.) *Hand-book for research on mathematics teaching and learning*. New York: MacMillan.
- Selwyn, N., Rivera-Vargas, P., Passeron, E., y Miño-Puigcerros, R. (2022). ¿Por qué no todo es (ni debe ser) digital? Interrogantes para pensar sobre digitalización, datificación e inteligencia artificial en educación. In P. Rivera-Vargas, R. Miño-Puigcerros, & E. Passeron (Eds.), *Educación con sentido transformador en la universidad* (pp. 137-147). Octaedro. <https://doi.org/10.31235/osf.io/vx4zr>
- Silva, M., Correa, R., y Mc-Guire, P. (2024). Metodologías Activas con Inteligencia Artificial y su relación con la enseñanza de la matemática en la educación superior en Chile. Estado del arte. *TE&ET*, 37, 20-29
- Stefanova, T., y Georgiev, S. (2024). Possibilities for using AI in mathematics education. *Mathematics and Education in Mathematics*. 53, 117–125. <https://doi.org/10.55630/mem.2024.53.117-125>.
- Sureda, P., Corica, A., Parra, V., Godoy, D., y Schiaffino, S. (2024). La evaluación en educación matemática: aportes de chatbots y futuros profesores de matemática. *EduTec*, (89), 64–83. <https://doi.org/10.21556/edutec.2024.89.3243>

- Vaillant, D., Rodríguez Zidán, E., y Bentancor Biagas, G. (2020). Uso de plataformas y herramientas digitales para la enseñanza de la Matemática. *Ensaio: Avaliação e Políticas Públicas em Educação*, 28(108). <https://doi.org/10.1590/S0104-40362020002802241>
- Velázquez, R., Merchán García, W., Maldonado Zúñiga, K., y Castro Landin, A. (2021). Metodología del aprendizaje basado en problemas aplicada en la enseñanza de las Matemáticas. *Serie Científica de la Universidad de las Ciencias Informáticas*, 14(3), 142-155.

Autora de correspondencia

ANA ROSA CORICA

Dirección: Facultad de Ciencias Exactas, Gral. Pinto 399, C.P. B7000GHG,
Tandil, Buenos Aires, Argentina.
acorica@niem.exa.unicen.edu.ar