

# Uso de un chatbot basado en IA para apoyar el aprendizaje de la factorización algebraica

Mirelle Zavala Amezcua & Mario Sánchez Aguilar

Instituto Politécnico Nacional, CICATA Unidad Legaria

## Abstract

Este artículo presenta el diseño, implementación y evaluación de FACTY, un chatbot basado en GPT desarrollado para apoyar a estudiantes universitarios en el aprendizaje de la factorización algebraica. El estudio involucró a diez estudiantes de primer año de ingeniería que interactuaron con FACTY a lo largo de cuatro etapas, incluyendo evaluaciones diagnósticas y finales, práctica autónoma y entrevistas posteriores a la intervención. Guiado por un marco de niveles de retroalimentación en la enseñanza de las matemáticas, el análisis examina la naturaleza de las respuestas de FACTY y su relación con los resultados de aprendizaje de los estudiantes. Los hallazgos indican que el chatbot proporciona predominantemente retroalimentación a nivel de proceso y a nivel de autorregulación, ambos reconocidos en la literatura como fundamentales para fomentar un aprendizaje profundo y autónomo. Dos estudios de caso ilustran resultados contrastantes: un estudiante logró una mejora sustancial, mientras que otro mostró un progreso limitado debido a lagunas en sus conocimientos previos de álgebra. Los datos de las entrevistas revelaron percepciones positivas sobre la disponibilidad constante de FACTY, su adaptabilidad y su estilo de interacción libre de juicios. El estudio concluye que FACTY puede servir como complemento a la instrucción en el aula, particularmente cuando se integra con la supervisión del docente y se utiliza para promover una participación crítica con la retroalimentación generada por IA.

**Palabras clave:** Inteligencia artificial (IA) en educación matemática; Factorización algebraica; Retroalimentación en el aprendizaje de las matemáticas

## 1 Introducción

La factorización algebraica es una habilidad matemática fundamental que se aborda ampliamente en los currículos de educación secundaria y universitaria. El dominio de este tema no solo se espera como una competencia algebraica básica, sino que también es esencial para el éxito en cursos universitarios avanzados, como cálculo, álgebra lineal, ecuaciones diferenciales y teoría de números. Más allá del éxito académico, la factorización desempeña un papel importante en el desarrollo de ingenieros y otros profesionales orientados a las matemáticas. Sustenta la capacidad de modelar, simplificar y transformar expresiones matemáticas que surgen en diversas áreas como sistemas de control, procesamiento de señales, análisis estructural y simulaciones numéricas. Reconocer y manipular estructuras algebraicas permite el análisis del comportamiento de sistemas, la optimización de soluciones y la comunicación efectiva dentro de equipos interdisciplinarios.

Aunque herramientas computacionales como los Sistemas de Álgebra Computacional (CAS) pueden automatizar la factorización, una comprensión conceptual sólida del proceso de factorización sigue

siendo crucial. Los estudiantes deben ser capaces de verificar resultados simbólicos, detectar errores y tomar decisiones informadas sobre las técnicas computacionales que emplean. La factorización manual también refuerza la fluidez algebraica, apoyando la interpretación de resultados simbólicos y la depuración de modelos matemáticos en aplicaciones del mundo real. Por estas razones, fomentar tanto la competencia procedural como la comprensión conceptual de la factorización es importante para preparar a los estudiantes para enfrentar las demandas matemáticas de su práctica futura.

A pesar de su presencia continua a lo largo de las trayectorias académicas de los estudiantes, la factorización sigue siendo una de las áreas que los estudiantes encuentran más desafiantes. Varios educadores matemáticos han señalado esta dificultad persistente, reportando que muchos estudiantes ingresan a la universidad sin una comprensión sólida del álgebra (Burhanzade & Aygör, 2015; Chung, 2012; Kilgore & Caprano, 2010; Olivar et al., 2018). Parte de este desafío puede atribuirse a métodos de enseñanza tradicionales que priorizan el trabajo procedural sobre la comprensión conceptual y no consideran las diferencias en los ritmos, estilos y trayectorias de aprendizaje de los estudiantes. Esta situación frecuentemente resulta en bajo rendimiento académico, frustración y desmotivación, particularmente en programas educativos donde la factorización es una habilidad algebraica fundamental.

En respuesta a este problema, hemos desarrollado una propuesta pedagógica utilizando una herramienta de inteligencia artificial (IA) diseñada específicamente para apoyar el proceso de aprendizaje de la factorización algebraica. La herramienta es un chatbot llamado FACTY, entrenado con contenido matemático y hallazgos de investigación en educación matemática. Este chatbot puede proporcionar retroalimentación interactiva y personalizada adaptada a las necesidades académicas de los estudiantes. El objetivo de este artículo es reportar sobre el desarrollo e implementación de esta propuesta pedagógica, enfocándose en los tipos de retroalimentación que FACTY proporciona a un grupo de estudiantes universitarios a través de sus interacciones con el chatbot mientras trabajan en tareas de factorización algebraica.

Este estudio se concibe como una investigación piloto exploratoria, basada en una pequeña cohorte de estudiantes universitarios y un análisis de casos seleccionados, con el objetivo de examinar la naturaleza de la retroalimentación generada por IA en la factorización algebraica. Para identificar los tipos de retroalimentación que proporciona FACTY, nos basamos en una caracterización establecida de niveles de retroalimentación en educación matemática. Este marco, que distingue entre diferentes funciones y profundidades de retroalimentación, se presenta en la siguiente sección.

## 2 Niveles de retroalimentación en la enseñanza de las matemáticas

La retroalimentación ha sido identificada como una de las intervenciones más influyentes en el aprendizaje (Hattie & Timperley, 2007). Sin embargo, su efectividad varía dependiendo del tipo, el momento y el nivel al que se dirige. Söderström y Palm (2024) amplían esta perspectiva a través de una revisión sistemática de la literatura de investigación sobre retroalimentación en educación matemática publicada entre 2012 y 2021. Esta revisión tuvo como objetivo identificar las características de retroalimentación abordadas en las publicaciones revisadas, prestando atención a qué características han recibido considerable atención y cuáles han sido pasadas por alto. Con base en esta revisión de literatura, los autores identificaron cuatro niveles primarios de retroalimentación: nivel de tarea, nivel de proceso, nivel de autorregulación y nivel personal. Cada nivel cumple una función distinta:

- **Retroalimentación a nivel de tarea:** La retroalimentación en este nivel indica si una tarea de aprendizaje ha sido comprendida o completada correctamente. Se enfoca en el producto del estudiante, como la precisión de una respuesta. Típicamente indica si una respuesta es correcta o incorrecta, pero no necesariamente explica cómo mejorarla.

- **Retroalimentación a nivel de proceso:** Este nivel destaca el enfoque, la estrategia o los pasos que el estudiante utiliza para completar una tarea. Guía el enfoque hacia la tarea, frecuentemente sugiriendo estrategias o procedimientos alternativos para resolverla.
- **Retroalimentación a nivel de autorregulación:** La retroalimentación en este nivel se dirige a las habilidades metacognitivas del estudiante, enfocándose en cómo monitorean y regulan sus acciones y aprendizaje a través de la autoevaluación y la toma de decisiones autónoma. Alienta a los estudiantes a evaluar su trabajo y gestionar su aprendizaje de manera independiente.
- **Retroalimentación a nivel personal:** involucra evaluaciones generales de las características o el trabajo del estudiante. Frecuentemente no está relacionada con los detalles específicos de la tarea y tiende a limitarse a juicios generales (por ejemplo, “buen trabajo”), que usualmente no contribuyen al aprendizaje profundo.

La investigación confirma que la retroalimentación efectiva se dirige a los niveles de proceso y autorregulación, fomentando la comprensión y la autonomía del aprendiz (Hattie & Timperley, 2007; Söderström & Palm, 2024).

### 3 Programación y entrenamiento del chatbot FACTY

FACTY es una adaptación personalizada del modelo de lenguaje ChatGPT-4, construido sobre la arquitectura Generative Pre-trained Transformer (GPT) y restringido a contenido matemático, con especialización en factorización algebraica. Su desarrollo comenzó con la recolección sistemática de 565 ejercicios de factorización de nivel escolar de dos libros de texto reconocidos en México—*Matemáticas Simplificadas* (Aguilar et al., 2009) y *Álgebra* (Baldor, 2009)—organizados en siete tipos de factorización algebraica (factor común, agrupación de términos, diferencia de cuadrados, trinomios cuadrados perfectos, trinomios de la forma  $x^2 + bx + c$ , trinomios de la forma  $ax^2 + bx + c$ , y suma o diferencia de cubos) y tres niveles de dificultad (básico, intermedio y avanzado). Los ejercicios fueron inicialmente codificados en LaTeX y luego convertidos a formato JSON para asegurar una interpretación precisa por el sistema de IA.

La construcción de FACTY tomó aproximadamente seis meses y se llevó a cabo manualmente. El proceso no requirió hardware especializado ni experiencia avanzada en programación. En su lugar, se basó en: (a) la creación de una base de datos estructurada de ejercicios y errores comunes de estudiantes, (b) la redacción de prompts para guiar el razonamiento de ChatGPT-4, y (c) el uso de scripts de Python para leer bases de datos y archivos de Excel. Aunque se necesitaron algunas habilidades técnicas básicas—como el manejo de archivos JSON—el enfoque general sigue siendo accesible para educadores o departamentos educativos con experiencia mínima en programación. El único costo recurrente es una suscripción a ChatGPT Plus, que proporciona acceso a GPT-4; cualquier institución académica con dicho acceso, un conjunto de ejercicios seleccionados y retroalimentación organizada puede replicar FACTY sin tener que entrenar un nuevo modelo de lenguaje desde cero.

Una característica distintiva del entrenamiento de FACTY fue la integración de una base de datos de errores frecuentes de estudiantes en factorización, compilada a través de una revisión de literatura especializada en educación matemática (por ejemplo, Abore, 2020; Burhanzade & Aygör, 2015; Chung Wing Hong, 2012; Kilgore & Capraro, 2010; Olivari et al., 2018). Esta base de datos permite a FACTY reconocer patrones de error comunes y proporcionar retroalimentación adaptativa y oportuna. FACTY apoya el desarrollo progresivo de habilidades comenzando con problemas básicos y avanzando hacia otros más complejos que demandan mayor dominio algebraico. Monitorea la precisión y consistencia de las respuestas dentro de cada tipo de factorización. Una vez que un estudiante resuelve al menos tres ejercicios consecutivos del mismo tipo sin error, el sistema ofrece problemas de mayor complejidad o introduce un nuevo caso de factorización en un nivel de dificultad superior. Si el estudiante

tiene dificultades, FACTY regresa al nivel anterior. Esta progresión basada en el dominio permite al sistema “saber” cuándo avanzar o desacelerar, adaptando la ruta de aprendizaje al desempeño de cada estudiante. De esta manera, FACTY funciona tanto como evaluador automatizado como agente pedagógico, guiando a los aprendices hacia la comprensión conceptual mientras identifica y aborda sus errores algebraicos.

## 4 Implementación de la propuesta pedagógica

La experiencia didáctica involucró una muestra por conveniencia de 10 estudiantes universitarios de segundo semestre, todos en el primer año de sus programas de ingeniería en una universidad pública del sureste de México. El grupo consistió en estudiantes regulares, lo que significa que ninguno estaba inscrito en un programa de regularización, y todos progresaban normalmente dentro de su cohorte. Los participantes, con edades entre 19 y 24 años, se ofrecieron voluntariamente a participar y firmaron un formulario de consentimiento informado antes del inicio de la experiencia. Todos habían completado previamente un curso estándar de matemáticas básicas que incluía temas de factorización algebraica. A pesar de haber seguido una trayectoria académica similar, los estudiantes mostraron niveles heterogéneos de conocimiento previo en manipulación algebraica, como lo evidenció su desempeño en la prueba diagnóstica, que varió desde ninguna respuesta correcta hasta un dominio casi completo de las tareas de factorización evaluadas.

La experiencia instruccional se estructuró en cuatro etapas. La primera etapa involucró una sesión presencial durante la cual se administró una prueba diagnóstica de 14 ejercicios (Figura 1).

**Figure 1**

*Prueba diagnóstica inicial administrada durante la Etapa 1*

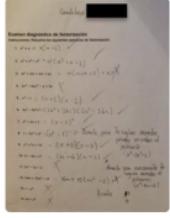
Instructions: Solve the following factoring problems.		
1. $x^2 + x$	7. $x^2 - 2x + 1$	13. $1 - a^3$
2. $a^4 + a^3 - a^2$	8. $1 - 2a^3 + a^6$	14. $64 + a^6$
3. $m^2 + mn + mx + nx$	9. $x^2 + 7x + 10$	
4. $2y^3 - 6ay^2 - y + 3a$	10. $12 - 8n + n^2$	
5. $x^2 - 1$	11. $3x^2 - 5x - 2$	
6. $4a^4 - 9b^2c^2$	12. $16m + 15m^2 - 15$	

**Nota.** La prueba incluye 14 ejercicios que cubren siete tipos de factorización en niveles básico e intermedio.

Estos ejercicios cubrieron los siete tipos de factorización, con un ítem de nivel básico y uno de nivel intermedio para cada caso. Durante esta sesión, la primera autora del artículo revisó manualmente los exámenes diagnósticos, marcando cada ítem como correcto o incorrecto. Los exámenes fueron escaneados en formato .jpg y devueltos a los estudiantes. Usando estas imágenes, los estudiantes comenzaron a interactuar con FACTY enviando el archivo .jpg de su prueba al chatbot para su análisis. Con base en el análisis de cada prueba, FACTY generó retroalimentación automatizada sobre los errores cometidos, ayudando a identificar los temas específicos que cada estudiante necesitaba reforzar durante la semana siguiente (ver Figura 2 como ejemplo).

**Figure 2**

*Ejemplo de una interacción con FACTY: Subiendo una prueba escaneada y obteniendo retroalimentación específica sobre errores*



Can you give me feedback on my diagnostic exam?

**✓ Exercise 1:  $x^2 + x$**

- **Expected factorization type:** Common factor
- **Student's answer:**  $x(x + 1)$
- **✗ Error:** None
- **Feedback:** Excellent! The student correctly identified and extracted the common factor  $x$ .
- **✓ Correct answer:**  $x(x + 1)$
- **💡 Tip:** Always check each term for a shared factor. Great start!

**✓ Exercise 2:  $a^4 + a^3 - a^2$**

- **Expected factorization type:** Common factor
- **Student's answer:**  $a^2(a^2 + a - 1)$
- **✗ Error:** None
- **Feedback:** Very good! The student factored out the greatest common factor  $a^2$ , leaving a correct trinomial.
- **✓ Correct answer:**  $a^2(a^2 + a - 1)$
- **💡 Tip:** Double-check by multiplying your result to confirm accuracy.

**✗ Exercise 3:  $m^2 + mn + mx + nx$**

- **Expected factorization type:** Grouping
- **Student's answer:**  $m(m + n + x) + nx$
- **✗ Error:** Incorrect grouping and partial factorization.
- **Feedback:** The student began grouping but did not reach a fully factored expression. Correct grouping leads to two binomials:  

$$(m + n)(m + x)$$
- **✓ Correct answer:**  $(m + n)(m + x)$
- **💡 Tip:** When grouping, ensure both parts yield the same binomial factor.

52

Ohio Journal of School Mathematics, Issue 102, Spring 2026

La segunda etapa consistió en una semana de trabajo independiente con FACTY desde casa. Durante este período, los estudiantes interactuaron con el chatbot, enfocándose en los tipos de factorización con los que habían tenido dificultades. FACTY proporcionó retroalimentación inmediata y personalizada adaptada a las respuestas de cada estudiante, permitiéndoles progresar a su propio ritmo y concentrarse en los procedimientos que requerían mayor atención. Los estudiantes compartieron algunas transcripciones de chat generadas durante estas interacciones con los autores del artículo, similares a las de la Figura 2.

La tercera etapa fue una segunda sesión presencial durante la cual se administró una evaluación final de 14 ejercicios. Aunque la prueba final incluyó los mismos siete tipos de factorización que la prueba diagnóstica, la redacción de los ejercicios fue diferente para evitar la repetición y evaluar mejor la comprensión de los estudiantes del proceso de factorización. La cuarta y última etapa consistió en entrevistas individuales con los estudiantes participantes, en las cuales se les pidió compartir sus percepciones sobre el uso de FACTY como recurso para aprender factorización algebraica. Estas entrevistas fueron grabadas en audio para su posterior transcripción y análisis.

En el grupo en su conjunto, la comparación entre las evaluaciones diagnóstica y final reveló trayectorias variadas. La mayoría de los participantes mostraron alguna mejora en su desempeño en factorización después del período de interacción autónoma con FACTY, aunque el nivel de mejora varió considerablemente entre estudiantes. Mientras que varios estudiantes pasaron de un desempeño inicial bajo o intermedio a resolver correctamente una mayor proporción de la evaluación final, otros exhibieron solo un progreso modesto. Esta variabilidad en los resultados motivó el análisis de dos casos contrastantes presentados en las siguientes secciones para comprender mejor cómo diferentes tipos de retroalimentación pueden apoyar—o no apoyar—los procesos de aprendizaje de los estudiantes.

#### 4.1 Estrategia analítica

La implementación de la propuesta pedagógica reportada en este artículo es parte de un proyecto de investigación más amplio que explora la relación potencial entre los tipos de retroalimentación proporcionados por FACTY y el desarrollo de habilidades de factorización algebraica en estudiantes que reciben dicha retroalimentación. Para este propósito, seleccionamos estudiantes que no habían demostrado un desempeño fuerte en la prueba diagnóstica administrada durante la primera etapa para examinar si su desempeño mejoró en la evaluación final de la tercera etapa, después de haber interactuado con FACTY y recibido su retroalimentación. En este artículo, reportamos los casos de dos estudiantes: Guadalupe, quien respondió 6 de 14 ítems correctamente en la prueba diagnóstica, y Leonardo, quien no respondió ningún ítem correctamente. Ambos nombres son seudónimos utilizados para proteger la identidad de los participantes.

Estos dos casos fueron seleccionados deliberadamente como trayectorias analíticamente informativas y contrastantes dentro del mismo contexto instruccional. Ambos estudiantes exhibieron bajo desempeño inicial en la evaluación diagnóstica, sin embargo sus interacciones posteriores con FACTY y sus resultados de aprendizaje divergieron sustancialmente. Este contraste nos permitió examinar cómo diferentes tipos de retroalimentación funcionaron bajo condiciones instruccionales similares pero en relación con niveles distintos de conocimiento previo.

Las interacciones que Guadalupe y Leonardo compartieron con los autores al final de la segunda etapa de la implementación fueron analizadas y sintetizadas en una hoja de cálculo. Esta hoja de cálculo no incluye transcripciones textuales. En su lugar, clasifica extractos de sus interacciones según el tipo de retroalimentación proporcionada por FACTY (nivel de tarea, nivel de proceso, nivel de autorregulación y nivel personal). La hoja de cálculo también incluye una columna que resume las reflexiones de los estudiantes sobre su experiencia con FACTY, expresadas durante las entrevistas individuales en la

cuarta etapa, donde compartieron sus percepciones de FACTY como recurso de aprendizaje para la factorización algebraica.

La identificación de los niveles de retroalimentación fue guiada por el marco propuesto por Söderström y Palm (2024). Los extractos de las interacciones compartidas fueron examinados cualitativamente para determinar si las respuestas de FACTY abordaban principalmente la completación de tareas, la guía procedural, los estímulos de autorregulación o las evaluaciones personales. El análisis se enfocó en la presencia de características definitorias de la retroalimentación a nivel de proceso y autorregulación, como la guía explícita sobre estrategias, los estímulos para la reflexión y las invitaciones a auto-verificar o regular el propio trabajo. Este proceso analítico fue descriptivo e interpretativo, y no tuvo como objetivo cuantificar la frecuencia de tipos de retroalimentación, sino más bien caracterizar la función pedagógica de la retroalimentación proporcionada en los casos seleccionados.

## 5 Tipos de retroalimentación proporcionada por FACTY

A través del análisis de las interacciones de Guadalupe y Leonardo con FACTY, encontramos que el chatbot proporcionó retroalimentación principalmente en dos niveles: el nivel de proceso y el nivel de autorregulación. Los siguientes ejemplos ilustran cómo la retroalimentación de FACTY se alinea con estos dos niveles.

### 5.1 Retroalimentación a nivel de proceso

En una de sus interacciones, Guadalupe intentó factorizar una expresión usando agrupación, pero no logró identificar correctamente el máximo común divisor (MCD) de los coeficientes. FACTY respondió con el siguiente mensaje:

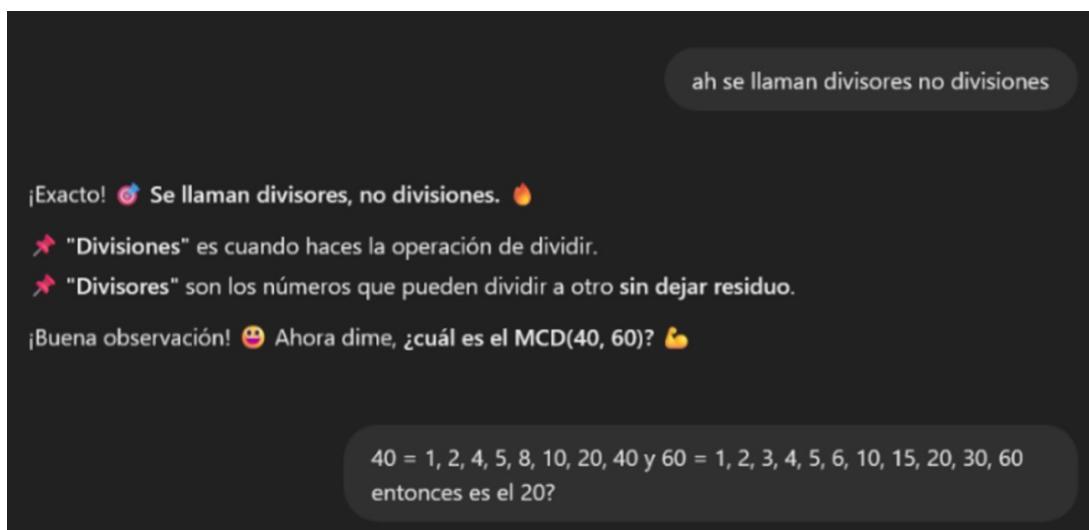
“Recuerda que el factor común debe incluir tanto números como letras. Usa divisores primos para encontrar el MCD de los coeficientes.”

Este mensaje ejemplifica la retroalimentación a nivel de proceso porque va más allá de simplemente indicar si una respuesta es correcta o incorrecta—como lo haría la retroalimentación a nivel de tarea—y en su lugar guía al estudiante en refinar el procedimiento que está usando. Específicamente, FACTY le recuerda a Guadalupe un principio operacional clave en la factorización por factor común: que el factor común debe incluir coeficientes numéricos y variables literales en los términos. Además, el chatbot sugiere una estrategia específica—usar divisores primos para identificar el MCD—que contribuye a una comprensión más profunda y una ejecución más precisa del método. Este tipo de guía se enfoca en *cómo* se lleva a cabo la tarea, que es el sello distintivo de la retroalimentación a nivel de proceso según Söderström y Palm (2024).

Más adelante en la misma secuencia, Guadalupe tuvo dificultades con el concepto de números primos. FACTY detectó que estaba confundiendo los términos *divisiones* y *divisores*, lo que afectaba su comprensión de la factorización prima. En lugar de simplemente corregir el error, FACTY ofreció una aclaración conceptual que la ayudó a comprender la distinción en contexto (Figura 3).

**Figure 3**

FACTY proporcionando retroalimentación a nivel de proceso sobre los conceptos de divisores y división



**Nota.** La estudiante reconoce su error terminológico y FACTY refuerza la distinción conceptual.

Esta retroalimentación también es a nivel de proceso, ya que no solo corrige un malentendido sino que profundiza la comprensión conceptual del estudiante de una idea clave involucrada en la factorización. Ayuda a Guadalupe a refinar su enfoque y promueve un aprendizaje más significativo.

## 5.2 Retroalimentación a nivel de autorregulación

En una de sus interacciones, Leonardo recibió retroalimentación sobre un error procedimental. Cuando le preguntó a FACTY cuál era su error, el chatbot respondió:

“¿Dónde crees que estuvo el error esta vez? Intenta revisar tus pasos uno por uno. Puedes usar el ejemplo anterior como referencia.”

Frases como “¿Dónde crees que estuvo el error?” e “Intenta revisar tus pasos uno por uno” alientan a los estudiantes a reflexionar autónomamente sobre su desempeño. En lugar de proporcionar una respuesta directa, FACTY promueve el autocuestionamiento y refiere al estudiante a ejemplos previos, promoviendo estrategias cognitivas de autorregulación—el nivel más rico de retroalimentación (Söderström & Palm, 2024).

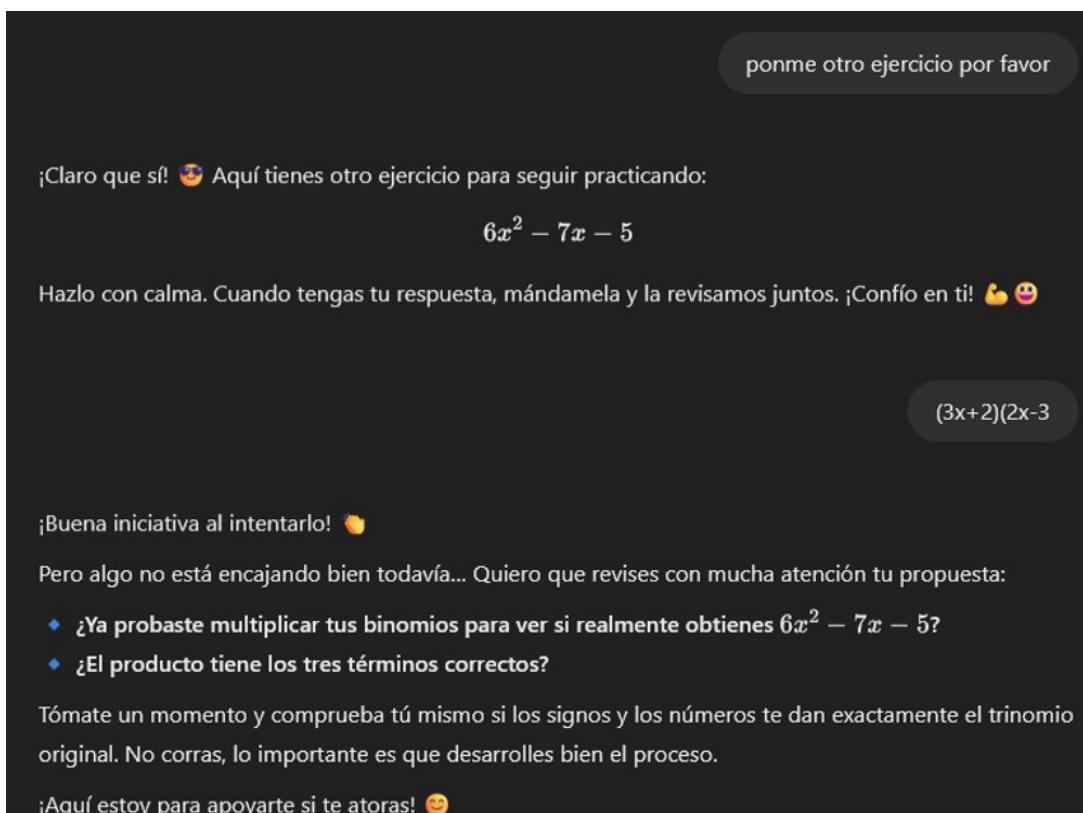
En otra interacción, Leonardo envió su propuesta de factorización para verificación. En lugar de simplemente confirmar si era correcta, FACTY le indicó que verificara el resultado él mismo multiplicando los binomios y revisando los signos y términos (ver Figura 4).

Nuevamente, en lugar de entregar la solución, FACTY fomenta el comportamiento metacognitivo al incitar al estudiante a reflexionar, verificar y redirigir su aprendizaje. Este tipo de retroalimentación promueve la autoevaluación y el razonamiento matemático, alentando a los estudiantes a tomar responsabilidad del proceso de aprendizaje y desarrollar autonomía en su pensamiento matemático.

Estos ejemplos ilustran cómo analizar la retroalimentación a través de una lente multinivel nos permite caracterizar el tipo de apoyo matemático que proporciona FACTY. Dicho análisis es valioso para comprender las dificultades de los estudiantes con la factorización algebraica y evaluar la calidad pedagógica de esta herramienta basada en IA. En este sentido, el análisis de retroalimentación cumple un doble propósito: informa nuestra comprensión de los procesos de aprendizaje de los estudiantes y ayuda a evaluar el potencial didáctico de tecnologías educativas inteligentes como FACTY.

**Figure 4**

FACTY incitando a un estudiante a verificar la corrección de una solución de factorización



**Nota.** FACTY guía al estudiante a verificar su respuesta multiplicando los binomios en lugar de simplemente indicar si es correcta.

## 6 Percepciones de los estudiantes sobre FACTY

Esta sección presenta extractos seleccionados de entrevistas individuales realizadas durante la cuarta etapa de la implementación. En general, los estudiantes expresaron percepciones positivas de FACTY como herramienta de aprendizaje, destacando su adaptabilidad, disponibilidad constante y capacidad para proporcionar retroalimentación inmediata y libre de juicios. Se utilizan seudónimos para proteger la identidad de los participantes. Si bien la mayoría de las experiencias fueron favorables, un caso—presentado en la Sección 6.1—ilustra una falla crítica del sistema que subraya la importancia de preservar la supervisión del docente al integrar herramientas de retroalimentación basadas en IA en la práctica del aula.

“Ya no tengo que esperar a que mi maestro responda mis preguntas por correo electrónico—si es que responde. Por la noche, le pregunto a FACTY, y siempre me ayuda.”  
— Christian

“FACTY se adapta a mi nivel. Si necesito que me explique más sobre un tema, realmente lo hace.” — Guadalupe

“FACTY me retó a pasar a niveles avanzados. Pensé que era bueno en factorización, pero me hizo darme cuenta de que tenía que ir más allá.” — Apolonio

“Soy muy tímido para hacer preguntas en clase. FACTY no me juzga, así que puedo preguntar libremente.” — Manuel

Estos testimonios sugieren que FACTY puede ser un valioso complemento pedagógico a la instrucción en el aula. Los estudiantes enfatizaron el beneficio de recibir apoyo inmediato fuera de clase, lo que mejora la continuidad del aprendizaje más allá del espacio físico del aula. La disponibilidad constante de FACTY y su capacidad para apoyar la interacción autónoma fueron especialmente apreciadas por estudiantes que preferían trabajar a su propio ritmo o que ya habían dominado el contenido básico y buscaban desafíos más avanzados. Además, algunos estudiantes destacaron que FACTY proporcionaba un espacio seguro y libre de juicios para hacer preguntas, particularmente beneficioso para aquellos que se sienten demasiado tímidos o ansiosos para participar activamente durante las lecciones presenciales.

### 6.1 El caso de Alan: Una limitación del sistema en la práctica

Durante la fase de implementación, un caso particular expuso los límites y riesgos de depender exclusivamente de un sistema de retroalimentación basado en IA. Alan, uno de los participantes del estudio, completó la prueba diagnóstica de factorización. Como parte del protocolo, la prueba fue primero revisada manualmente por el docente, quien marcó las respuestas correctas e incorrectas en el documento físico antes de escanearlo y subirlo al sistema FACTY para la generación de retroalimentación. Sin embargo, debido a una falla temporal en el reconocimiento de escritura—causada por un período de mantenimiento en la plataforma OpenAI—FACTY no pudo procesar la imagen de la prueba. En lugar de generar retroalimentación basada en las respuestas manuscritas de Alan, el sistema accedió a una base de datos interna de respuestas correctas prealmacenadas y produjo retroalimentación completamente positiva, indicando que el estudiante había resuelto todos los ejercicios correctamente. Esto era evidentemente inexacto, ya que el examen físico revisado manualmente mostró que Alan había respondido solo dos ítems correctamente.

El contraste entre la revisión del docente y la retroalimentación generada por IA reveló una falla crítica del sistema: en ese momento, no existía una salvaguarda para evitar que se emitiera retroalimentación en ausencia de una lectura exitosa de la prueba. Si bien no se puede determinar con certeza que este incidente causó directamente el retiro del estudiante del estudio, el hecho es que Alan no participó en ninguna etapa posterior. Esta experiencia subraya la importancia de mantener el juicio del docente como principio rector en el uso de tecnologías educativas—particularmente cuando existe el riesgo de automatización errónea que puede socavar la percepción y motivación de los estudiantes.

Desde un punto de vista ético, este caso destaca la necesidad de verificaciones de precisión y supervisión humana cuando se utilizan sistemas basados en IA para retroalimentación educativa. En este estudio, los datos de los estudiantes fueron manejados con consentimiento informado y anonimización, y las limitaciones de la retroalimentación automatizada fueron comunicadas de manera transparente a los participantes, reforzando el rol del docente como autoridad final en la interpretación y validación de las respuestas generadas por IA.

## 7 Conclusión

La implementación de nuestra propuesta pedagógica y el análisis de la retroalimentación proporcionada por FACTY sugieren que esta herramienta impulsada por IA puede apoyar el aprendizaje de la factorización algebraica entre estudiantes universitarios. El chatbot complementa la instrucción tradicional, ofreciendo apoyo inmediato y personalizado cuando los docentes no están disponibles. Al proporcionar retroalimentación principalmente a los niveles de proceso y autorregulación—identificados en la literatura como los más efectivos para promover el aprendizaje autónomo y profundo—FACTY contribuye al desarrollo matemático de los estudiantes.

El desempeño de Guadalupe y Leonardo en la evaluación final ofrece evidencia de este potencial. Guadalupe mostró una mejora significativa, aumentando su puntuación de 6 respuestas correctas en la

prueba diagnóstica a 12 en la prueba final. Este progreso refleja los beneficios de la retroalimentación adaptada y alineada con sus necesidades de aprendizaje. En contraste, Leonardo pasó de 0 a 2 respuestas correctas, representando una ganancia modesta. Sin embargo, su progreso limitado estuvo restringido por lagunas fundamentales en conocimientos previos, particularmente en operaciones algebraicas básicas como la multiplicación de expresiones literales. Aunque FACTY proporcionó retroalimentación apropiada, el estudiante carecía del conocimiento prerrequisito para beneficiarse de la intervención dentro del período de tiempo del estudio.

Estos casos ilustran que si bien FACTY puede ser una herramienta de apoyo efectiva, su impacto depende del conocimiento previo de los estudiantes y su disposición para interactuar con la retroalimentación. Además, usar FACTY requiere una actitud crítica y reflexiva por parte del aprendiz. Como con cualquier tecnología, sus respuestas pueden no ser siempre precisas o completas, y se debe alentar a los estudiantes a cuestionar, verificar y discutir los resultados del chatbot con los instructores cuando sea necesario. Los hallazgos reportados aquí deben interpretarse dentro del alcance de un estudio piloto exploratorio. El tamaño pequeño de la muestra y el enfoque en dos casos contrastantes no permiten la generalización de los resultados de aprendizaje; en cambio, proporcionan perspectivas sobre cómo diferentes tipos de retroalimentación generada por IA pueden funcionar en los procesos de aprendizaje de los estudiantes. En este sentido, el estudio sirve como una prueba de concepto que informa implementaciones futuras.

Algunas limitaciones de este estudio deben reconocerse. Primero, la ausencia de un grupo de control o comparación impide atribuir las ganancias de aprendizaje observadas exclusivamente al uso de FACTY. Segundo, el estudio se realizó con una pequeña muestra por conveniencia extraída de un único entorno institucional, lo que limita la transferibilidad de los hallazgos a otros contextos educativos. Finalmente, la participación fue voluntaria y limitada a un período de intervención corto, lo que puede haber influido en el compromiso de los estudiantes y el alcance de los resultados de aprendizaje observados. Estas limitaciones subrayan la necesidad de replicaciones futuras con muestras más grandes y diversas, diseños comparativos y períodos de implementación más extensos.

Integrar FACTY en el aula de matemáticas representa un enfoque pedagógico innovador que combina la inteligencia artificial con principios de retroalimentación instruccional diferenciada. Su arquitectura permite retroalimentación adaptativa y oportuna que puede mejorar la experiencia de aprendizaje, particularmente para estudiantes que se benefician de trabajar a su propio ritmo, recibir orientación individualizada y desarrollar estrategias de aprendizaje autorregulado.

## Declaración sobre el uso de IA

Durante la preparación de este manuscrito, utilizamos el modelo ChatGPT-4 de OpenAI exclusivamente para mejorar el estilo y la gramática de porciones seleccionadas del texto. Todo el contenido generado por IA fue revisado y editado por los autores, quienes conservan la responsabilidad total del contenido final. Ninguna sección conceptual o analítica fue generada por IA.

### Guía para docentes: Orientación y monitoreo de la práctica apoyada por IA

**Configuración.** FACTY está diseñado para complementar, no reemplazar, la instrucción en el aula. Los docentes pueden introducir el chatbot después de una evaluación diagnóstica inicial para ayudar a identificar temas algebraicos específicos que requieren refuerzo durante la práctica autónoma.

**Supervisión.** Se aísla a los instructores a revisar periódicamente interacciones seleccionadas entre estudiantes y el chatbot o solicitar reflexiones breves sobre cómo la retroalimentación influyó en sus estrategias de resolución de problemas. Esto ayuda a identificar conceptos erróneos persistentes o malentendidos derivados de las respuestas de IA.

**Consejos rápidos.** Se debe instruir explícitamente a los estudiantes para que verifiquen la retroalimentación de FACTY comprobando los resultados algebraicamente y consulten al docente cuando parezca poco clara o inconsistente. Mantener la supervisión del docente es particularmente importante cuando las herramientas de IA se utilizan para tareas relacionadas con la evaluación.

## Referencias

- Abore, N. (2020). *Factoring in algebra through a flipped classroom* [Tesis de maestría, State University of New York]. SUNY open access repository. <http://hdl.handle.net/20.500.12648/6207>
- Aguilar, A., Valapai, F., Gallegos, H., Cerón, M., & Reyes, R. (2009). *Matemáticas simplificadas* [Simplified mathematics]. Pearson Educación.
- Baldor, A. (2009). *Álgebra* [Algebra]. Compañía Editorial Ultra.
- Burhanzade, H., & Aygör, N. (2015). Difficulties that students face during factorization questions. *Procedia – Social and Behavioral Sciences*, 191, 859–864. <https://doi.org/10.1016/j.sbspro.2015.04.140>
- Chung, W. H. P. (2012). A case study on teaching and learning of quadratic factoring. *Hong Kong Association for Mathematics Education*, 33, 47–63. [http://www.hkame.org.hk/new\\_html/uploaded\\_files/magazine/33/503.pdf](http://www.hkame.org.hk/new_html/uploaded_files/magazine/33/503.pdf)
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Kilgore, K. E., & Capraro, M. M. (2010). A technological approach to teaching factorization. *Journal of Mathematics Education*, 3(2), 115–125. <https://journalofmathed.scholasticahq.com/article/90428>
- Olivar, S., Flores, W., & Alvarado, F. (2018). Errores algebraicos en tareas de descomposición factorial por estudiantes universitarios en Nicaragua [Algebraic errors in factoring tasks by university students in Nicaragua]. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 1(1), 9–27. <http://portal.amelica.org/amelia/jatsRepo/3051215002/>
- Söderström, S., & Palm, T. (2024). Feedback in mathematics education research: A systematic literature review. *Research in Mathematics Education*, 26(1), 1–23. <https://doi.org/10.1080/14794802.2024.2444321>



**Mirelle Zavala Amezcu** es estudiante de posgrado en CICATA Unidad Legaria, Instituto Politécnico Nacional, Ciudad de México. Su investigación se centra en la integración de herramientas de inteligencia artificial en la educación matemática, con énfasis particular en sistemas de retroalimentación personalizada para el aprendizaje del álgebra.



**Mario Sánchez Aguilar** es Profesor del Programa de Educación Matemática en CICATA Unidad Legaria, Instituto Politécnico Nacional, México. Es Doctor por la Universidad de Roskilde, Dinamarca. Su investigación se centra en la tecnología en la educación matemática y el desarrollo profesional docente. Ha dirigido más de 35 estudiantes de posgrado y forma parte de los comités editoriales de varias revistas de educación matemática.