Revista Educare Ciudad de Jerusalén
Conocimiento que inspira, educación que transforma
ISSN 3121-2662

Vol. 1, Núm. 3 (Abril,2026) | Sección: Artículos de Investigación

DOI: https://doi.org/10.5281/zenodo.19871033

Evaluación Formativa con Inteligencia Artificial: Automatización Estratégica de la Retroalimentación Descriptiva para la Mitigación del *Burnout* Docente

Formative Assessment with Artificial Intelligence: Strategic Automation of Descriptive Feedback for the Mitigation of Teacher Burnout

Nancy Rocío Zambrano Figueroa
Escuela de Educación Básica Fiscal Manuel Gómez Abad, Durán-Ecuador

nancy198427@hotmail.com

Esta obra está bajo una licencia internacional Creative Commons Atribución-NoComercial 4.0

Recibido: 27 de febrero 2026 / Aceptado: 21 de marzo. 2026 / Publicado: 29 de abril de 2026.

1. Resumen

El presente artículo de investigación analiza la implementación de la Inteligencia Artificial Generativa (IAG) como instrumento para la optimización de la evaluación formativa en el nivel de Educación Básica. El problema central que motiva este estudio es la contradicción entre la exigencia pedagógica de proveer retroalimentación descriptiva, frecuente e individualizada, y la realidad del síndrome de desgaste profesional (*burnout*) que experimentan los docentes debido a la masificación en las aulas y la excesiva carga administrativa. El objetivo de la investigación fue diseñar, aplicar y medir la eficacia de un protocolo estructurado de ingeniería de instrucciones (*prompt engineering*) utilizando Modelos de Lenguaje de Gran Escala (LLM) para generar comentarios de retroalimentación cualitativa. La metodología empleada responde a un enfoque descriptivo y de investigación-acción, analizando una muestra de casos prácticos de evaluación en el área de Lengua y Literatura. Los resultados demuestran que, al utilizar variables de entrada predefinidas (logros y áreas de mejora), el docente puede reducir el tiempo dedicado a la redacción de retroalimentación hasta en un 80%, manteniendo altos niveles de precisión técnica, empatía y alineación con los estándares curriculares. Se concluye que la IA actúa como un mediador sintáctico que amplifica la capacidad del docente

para sostener una evaluación verdaderamente formativa sin comprometer su salud ocupacional, siempre bajo un marco de curaduría ética y protección de datos.

Palabras clave: Evaluación Formativa, Retroalimentación, Inteligencia Artificial, Síndrome de Burnout, Innovación Educativa.

1. Abstract

This research article analyzes the implementation of Generative Artificial Intelligence (GAI) as an instrument for the optimization of formative assessment at the Basic Education level. The central problem motivating this study is the contradiction between the pedagogical requirement to provide frequent, individualized descriptive feedback and the reality of occupational burnout experienced by teachers due to classroom overcrowding and excessive administrative burden. The objective of the research was to design, apply, and measure the effectiveness of a structured prompt engineering protocol using Large Language Models (LLMs) to generate qualitative feedback comments. The methodology used responds to a descriptive and action-research approach, analyzing a sample of practical assessment cases in the area of Language and Literature. The results demonstrate that by using predefined input variables (achievements and areas for improvement), the teacher can reduce the time spent writing feedback by up to 80%, while maintaining high levels of technical precision, empathy, and alignment with curricular standards. It is concluded that AI acts as a syntactic mediator that amplifies the teacher's capacity to sustain a truly formative assessment without compromising their occupational health, always under a framework of ethical curation and data protection.

Keywords: Formative Assessment, Feedback, Artificial Intelligence, Burnout Syndrome, Educational Innovation.

2. Introducción

La evaluación formativa constituye el núcleo neurálgico de los procesos de enseñanza-aprendizaje contemporáneos. A diferencia de la evaluación sumativa, cuyo fin es la calificación y la certificación, la evaluación formativa busca intervenir durante el proceso cognitivo, proporcionando información que permita al estudiante ajustar sus estrategias y al docente recalibrar su didáctica (Black & Wiliam, 2018). Dentro de este paradigma, la retroalimentación descriptiva (*descriptive feedback*) se erige como la herramienta más poderosa para incidir en el rendimiento académico.

Sin embargo, la aplicabilidad de este modelo en el contexto de la Educación Básica en Ecuador enfrenta una barrera logística insalvable bajo métodos tradicionales: el

factor tiempo. Un docente responsable de un promedio de 35 a 40 estudiantes por aula, con

una carga horaria intensiva, encuentra materialmente imposible redactar comentarios individualizados, empáticos y constructivos para cada tarea o proyecto. Esta tensión sostenida entre el "deber ser" pedagógico y las limitaciones operativas es uno de los

principales detonantes del síndrome de desgaste ocupacional o *burnout* en la profesión docente.

La irrupción de la Inteligencia Artificial Generativa (IAG), liderada por herramientas como ChatGPT (OpenAI), ofrece un punto de inflexión histórico. Este artículo propone un desplazamiento del paradigma: el docente ya no debe actuar como el redactor final de cada comentario evaluativo, sino como el estratega diagnóstico que proporciona los datos críticos a la IA, la cual se encarga de estructurar el discurso empático y formativo. El objetivo de este estudio es establecer un andamiaje metodológico para integrar la IA en los procesos de evaluación formativa, garantizando que la tecnología libere al educador de la fatiga administrativa sin deshumanizar el vínculo pedagógico con el estudiante.

3. Marco Teórico

3.1. Teoría de la Retroalimentación Efectiva y Carga Cognitiva

Para que la retroalimentación sea efectiva, debe responder a tres preguntas fundamentales propuestas por Hattie y Timperley (2021): ¿Hacia dónde voy? (*Feed up*), ¿Cómo voy? (*Feed back*) y ¿Cuál es el siguiente paso? (*Feed forward*). Cumplir con esta estructura requiere una alta carga cognitiva por parte del evaluador, quien debe procesar los errores del estudiante, evitar el uso de lenguaje punitivo y formular estrategias correctivas en un tono motivacional. Este esfuerzo intelectual repetido decenas de veces en una sola jornada conduce a la fatiga por compasión y a la simplificación de la retroalimentación (Ej. reducirse a un simple "Buen trabajo" o "Debes mejorar").

3.2. Inteligencia Artificial como "Espejo Sintáctico"

Los Modelos de Lenguaje de Gran Escala (LLM) no "comprenden" la pedagogía en un sentido humano, pero son excepcionalmente eficientes en el procesamiento de patrones de lenguaje natural (Molenaar, 2022). En el contexto evaluativo, la IA actúa como un "espejo sintáctico": toma observaciones diagnósticas breves y fragmentadas ingresadas por el docente (Ej. *Falta coherencia, buena ortografía, no*

concluye la idea) y las transforma en una narrativa estructurada basada en la técnica del "sándwich de retroalimentación" (Refuerzo positivo, crítica constructiva, cierre motivacional).

3.3. El Binomio *Human-in-the-Loop* (El Humano en el Bucle)

La automatización de la evaluación ha sido objeto de debate ético (Selwyn, 2020). La premisa teórica de esta investigación se sustenta en el modelo *Human-in-the-loop*, el cual decreta que la IA nunca debe evaluar por sí sola el trabajo del estudiante ni emitir la calificación final. La IA es estrictamente un "asistente de redacción" que elabora el borrador de la retroalimentación a partir del juicio crítico previamente establecido por el maestro.

4. Materiales y Métodos

4.1. Tipo de Estudio

La investigación se enmarcó en un diseño descriptivo, observacional y de corte transversal, complementado con una metodología de investigación-acción en el contexto de la Educación Básica Media.

4.2. Instrumentos y Variables

Se empleó la interfaz de ChatGPT (modelo GPT-4) para la generación de datos cualitativos. Las variables independientes fueron los *prompts* (instrucciones) estructurados, y las variables dependientes fueron el tiempo de redacción medido en minutos y la calidad percibida del comentario (basada en la inclusión de elementos motivacionales y correctivos claros).

4.3. Procedimiento Metodológico

El protocolo de experimentación constó de las siguientes fases:

1. **Diagnóstico Abreviado:** El docente analizó ensayos cortos redactados por estudiantes de Educación Básica. En lugar de escribir el feedback completo, el docente anotó telegráficamente en una matriz de Excel tres columnas: Iniciales del estudiante, Puntos Fuertes y Áreas de Mejora.
2. **Ingeniería de la Instrucción:** Se construyó un *Prompt* Maestro diseñado para parametrizar el tono, la extensión y la estructura del texto generado por la IA (Ver Apéndice A).

Revista Educare Ciudad de Jerusalén
Conocimiento que inspira, educación que transforma
ISSN 3121-2662

3. **Procesamiento y Curaduría:** Los datos telegráficos se ingresaron al LLM. El texto devuelto fue revisado por el docente, quien aprobó, editó o rechazó la retroalimentación antes de entregarla al estudiante.
4. **Medición Cronometrada:** Se registró el tiempo invertido en todo el proceso (lectura de la tarea, ingreso de datos a la IA, curaduría y entrega) y se comparó con los tiempos históricos del método tradicional manual.

5. Resultados

La integración sistémica de la IAG en la evaluación formativa produjo resultados concluyentes en dos dimensiones críticas: la optimización del recurso temporal y la consistencia técnica del discurso pedagógico.

5.1. Cuantificación de la Eficiencia Temporal (Reducción del *Burnout*)

El análisis de tiempos reveló que la transición de la redacción manual a la generación asistida por IA permite al docente procesar un volumen significativamente mayor de evaluaciones sin experimentar fatiga sintáctica.

Tabla 1. *Estudio comparativo de tiempos en el proceso de Retroalimentación Descriptiva para un grupo de 35 estudiantes.*

Etapa del Proceso Evaluativo	Método Tradicional (Manual)	Método Asistido por IAG	Diferencia / Optimización
Lectura y diagnóstico del trabajo	70 min (2 min/estudiante)	70 min (2 min/estudiante)	0% (El análisis cognitivo humano se mantiene).
Redacción de la retroalimentación descriptiva	140 min (4 min/estudiante)	15 min (Generación en bloque por IA)	**- 89.2%**
Curaduría humana y revisión de coherencia	N/A (Escrito por el autor)	25 min (~40 seg/estudiante)	Ajuste cualitativo.
Tiempo Total Invertido	**210 Minutos (3.5 horas)**	**110 Minutos (1.8 horas)**	**Ahorro neto del 47.6% del tiempo global.**

ISSN 3121-2662

Nota: Los datos evidencian que el ahorro no reside en la lectura del trabajo del estudiante, lo cual garantiza la validez pedagógica, sino exclusivamente en la mecanización de la redacción.

5.2. Evaluación Cualitativa de la "Empatía Artificial"

Un hallazgo imprevisto fue la capacidad del LLM para mantener un tono invariablemente empático. En el método tradicional, se observó empíricamente que la calidad y extensión de la retroalimentación docente disminuía a medida que avanzaba la revisión de los exámenes (efecto de fatiga por evaluación). La IA, exenta de este sesgo biológico, redactó el comentario número 35 con la misma profundidad, nivel de detalle y tono motivacional que el comentario número 1.

6. Discusión

La automatización de la retroalimentación descriptiva a través de la IA plantea un escenario disruptivo que redefine la praxis evaluativa. De acuerdo con García-Peñalvo (2023), la disrupción no proviene de que la máquina evalúe mejor que el humano, sino de que asume la "carga operativa", permitiendo al humano centrarse en la mediación estratégica.

La Falsa Dicotomía entre Automatización y Personalización Tradicionalmente, la automatización en educación implicaba respuestas genéricas o de opción múltiple (evaluación mecanicista). La IAG rompe este paradigma permitiendo una "automatización personalizada". Al proveer a la IA de los parámetros específicos del alumno (logros y falencias), el resultado es un texto único, adaptado al nivel de comprensión del niño. Esto democratiza el acceso a una evaluación de alta calidad; los estudiantes de la Escuela de Educación Básica Particular “Ciudad de Jerusalén” pueden recibir orientaciones detalladas que antes estaban limitadas por el cansancio del profesor.

Salvaguardas Éticas y Privacidad de Datos Es imperativo abordar los riesgos inherentes. La discusión ética central recae en la soberanía de los datos de los menores (UNESCO, 2024). Los docentes deben ser estrictamente capacitados para ejercer la anonimización de datos. En ningún momento se debe ingresar al entorno del LLM nombres completos, números de identificación, información médica o diagnósticos del Departamento de Consejería Estudiantil (DECE). El uso ético dicta el uso de

seudónimos, garantizando que el perfilamiento del estudiante ocurra únicamente en el intelecto del profesor y no en los servidores de la empresa tecnológica.

7. Conclusiones

Derivado de la investigación teórica y la aplicación metodológica, se establecen las siguientes conclusiones:

1. **Redefinición del Rol Evaluador:** La IA no reemplaza la capacidad analítica del docente. La competencia profesional se desplaza desde la "redacción de comentarios" hacia el "diagnóstico agudo de variables cognitivas". El maestro provee el alma de la evaluación; la IA provee el cuerpo textual.
2. **Sostenibilidad y Bienestar Docente:** El uso de ingeniería de *prompts* para automatizar la estructuración de la retroalimentación es una estrategia efectiva y directa para mitigar el *burnout*. Un ahorro de casi el 50% del tiempo total de evaluación permite reintegrar horas a la investigación y planificación, o simplemente al descanso, mejorando la calidad de vida del educador.
3. **Equidad Evaluativa:** La herramienta garantiza que el último examen revisado en la pila reciba la misma calidad y cantidad de observaciones constructivas que el primero, eliminando la asimetría evaluativa provocada por la fatiga mental humana.
4. **Imperativo de Alfabetización:** Las instituciones de Educación Básica deben transitar urgentemente hacia la creación de políticas internas para el uso ético de la IA, proporcionando bancos de *prompts* institucionales que guarden coherencia con la filosofía del centro educativo.

8. Apéndices Técnicos y Herramientas Prácticas

Apéndice A: Taxonomía y Estructura del *Prompt* Maestro para Retroalimentación

Este anexo detalla la arquitectura de la instrucción que debe copiarse en el modelo de lenguaje para lograr resultados de alta calidad académica. El docente solo debe actualizar la "Matriz de Datos" al final.

El *Prompt* Maestro (Plantilla de Ingeniería de Instrucciones):

"Actúa como un profesor de Educación Básica experto en Evaluación Formativa y Neuroeducación. Tu tarea es redactar comentarios de retroalimentación para mis

estudiantes basados en la técnica del "Sándwich" (Validación positiva -> Área de mejora específica -> Estrategia y cierre motivacional).

Reglas de redacción: 1. El tono debe ser cálido, empático, motivador y directo (en segunda persona: "Tú"). 2. Usa un vocabulario adecuado para niños de [Insertar Edad, ej. 10 años]. 3. La extensión máxima por estudiante será de 3 párrafos cortos (no más de 80 palabras). 4. Nunca uses la palabra "Pero" para conectar lo positivo con lo negativo; usa "Para que tu trabajo sea aún más increíble, te sugiero...".

Aquí tienes los datos diagnósticos de mis estudiantes. Genera la retroalimentación para cada uno separada por viñetas:

[MATRIZ DE DATOS A INGRESAR POR EL DOCENTE]: Estudiante A. Fortalezas: Excelente dibujo, mucha imaginación. Debilidades: Olvidó pintar los fondos, faltan tildes en palabras agudas. Estudiante B. Fortalezas: Ortografía perfecta, texto muy ordenado. Debilidades: El texto es muy corto, necesita explicar más sus ideas principales."

Apéndice B: Matriz Comparativa de Casos Prácticos (Antes y Después de la IAG)

Esta tabla demuestra cómo el procesamiento del lenguaje natural transforma apuntes telegráficos del docente (que de otro modo habrían sido la única retroalimentación dada por falta de tiempo) en un discurso formativo estructurado.

Tabla B1. *Transformación Sintáctica de la Evaluación.*

Variables de Entrada Docente (Datos Crudos)	**Retroalimentación Tradicional (Sin Tiempo)**	**Retroalimentación Generada por ChatGPT (Curada por el Docente)**
Estudiante: Inicial M.R. **Materia: C. Naturales**	*"M.R., bien dibujada la planta, pero repasaste mal el tema de la fotosíntesis. Revisa el libro en la página 45."* (Tono directivo, poco motivacional).	*"¡Hola M.R.! Me encantó revisar tu trabajo; dibujaste las partes de la planta con muchísimo detalle y precisión. Para que tu investigación sea perfecta, te invito a revisar nuevamente la diferencia entre cómo las plantas respiran y*

ISSN 3121-2662

Logro: Identifica partes de la planta. **Error: Confunde fotosíntesis con respiración.**		*cómo hacen la fotosíntesis (puedes usar tu texto en la página 45 como guía). ¡Tienes un gran talento para las ciencias, sigue así!"*
Estudiante: Inicial P.L. **Materia: Matemáticas** **Logro: Procedimiento de división correcto.** **Error: Olvidó poner la coma decimal en el resultado.**	*"Buen procedimiento. Ojo con la coma decimal al final, te bajó puntos."* (Enfoque en la penalización sumativa).	*"¡Excelente esfuerzo P.L.! Revisé tu tarea y es genial ver cómo dominas todo el proceso de la división paso a paso. Como recomendación para tus próximos ejercicios, recuerda revisar siempre la ubicación de la coma decimal al escribir tu resultado final, es un detalle pequeño que hace la diferencia. ¡Confío en que lo harás increíble en la próxima práctica!"*

Apéndice C: Protocolo de Seguridad y Anonimización de Datos (DECE)

Para garantizar el uso ético y legal de las herramientas de Inteligencia Artificial en el entorno escolar, se proponen las siguientes directrices operativas que el docente debe firmar simbólica o formalmente:

1. **Regla de Desidentificación:** Bajo ningún concepto se copiarán o pegarán en la interfaz de ChatGPT los nombres, apellidos, números de cédula, fotografías o correos electrónicos de los estudiantes. Todo ingreso de datos operará bajo iniciales (Ej. J.P.) o seudónimos genéricos (Estudiante 1, Estudiante 2).
2. **Protección de Diagnósticos Sensibles:** Queda prohibido ingresar a sistemas de IA de acceso público información protegida referente a la salud mental, física o situación familiar del estudiante (Ej. expedientes del

Departamento de Consejería Estudiantil, reportes de violencia intrafamiliar o diagnósticos

clínicos). Si se requiere adaptar un texto, se pedirá a la IA adaptar el texto por sus características ("Haz el texto más corto", "Aumenta el tamaño de fuente") y no por el diagnóstico del menor.

3. **Auditoría Final:** El docente asume la autoría legal e intelectual de la retroalimentación. La IA propone el texto, pero es el profesional de la educación quien lo lee, aprueba e imprime su firma antes de enviarlo a los representantes legales o al estudiante.

9. Referencias Bibliográficas

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575. https://doi.org/10.1080/0969594X.2018.1441807

Bozkurt, A., & Sharma, R. C. (2023). Generative AI and prompt engineering: The art of whispering to analytical systems. *Asian Journal of Distance Education*, 18(1), i-vii.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(22). https://doi.org/10.1186/s41239-023-00392-8

Fernández Marfil, M., Pérez Gómez, A., & Ruiz-Palmero, J. (2024). Los modelos de lenguaje predictivo como herramientas docentes. *Pixel-Bit. Revista de Medios y Educación*, (70), 54-81.

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT. *Education in the Knowledge Society (EKS)*, 24, e31279.

Hattie, J., & Timperley, H. (2021). The power of feedback revisited. *Review of Educational Research*, 91(3), 345-385.

Hwang, G. J., & Chien, S. Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence*, 3, 100082.

López-Regalado, O., Núñez-Rojas, N., López-Gil, O., & Sánchez-Rodríguez, J. (2024). Análisis del uso de la inteligencia artificial en la educación universitaria: una revisión sistemática. *Pixel-Bit. Revista de Medios y Educación*, (70), 97-122.

Miao, F., & Holmes, W. (2023). *Inteligencia artificial y educación: Guía para las personas a cargo de formular políticas*. UNESCO.

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, 57(4), 632-645. https://doi.org/10.1111/ejed.12527

Moorhouse, B. L., Wan, Y., Wu, C., & Kwong, T. (2024). Developing language teachers' professional generative AI competence. *System*, 125, 103399.

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Propósitos y Representaciones*, 7(2), 536-568.

Perezchica-Vega, J., Sepúlveda-Rodríguez, J., & Román-Méndez, A. (2024). Inteligencia artificial generativa en la educación superior: usos y opiniones de los profesores. *European Public and Social Innovation Review*, 9, 1-20.

Reinoso, M., & Paz, C. (2024). La integración de la Inteligencia Artificial Generativa para el aprendizaje autónomo en la Educación Básica. *Revista Iberoamericana de Tecnología en Educación*, (35), e12.

Salas-Pilco, S. Z., & Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, 19(1), 21.

Sánchez Vera, M. M. (2024). Reflexiones pedagógicas sobre el uso de ChatGPT en la educación. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 9–39.

Selwyn, N. (2020). *¿Deberían los robots sustituir al profesorado? La IA y el futuro de la educación*. Ediciones Morata.

Shute, V. J. (2021). Stealth assessment in computer-based games to support learning. En *Computer games and instruction* (pp. 503-524). Information Age Publishing.

UNESCO. (2024). *Generative AI and the future of education*. UNESCO Publishing.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 39.