

## IA generativa y ChatGPT en Educación: Un reto para la evaluación y ¿una nueva pedagogía?

### Generative AI and ChatGPT in Education: A challenge for evaluation and a new pedagogy?

Miguel Zapata Ros

Revista de Educación a Distancia (RED), España

<https://orcid.org/0000-0003-4185-5024>

E-mail: [mzapata@um.es](mailto:mzapata@um.es)

#### Resumen

En el presente artículo se aborda la influencia de la inteligencia artificial (IA) generativa, específicamente el modelo ChatGPT, en el ámbito educativo. Se inicia con la exploración de la naturaleza de los programas generativos, como los modelos Transformer AI, y su impacto en la educación. Se examina la pregunta fundamental sobre qué son estos programas y cómo utilizan estadísticas de grandes conjuntos de datos para generar texto. El artículo destaca tanto las repercusiones positivas como las negativas de la presencia de ChatGPT en la educación. En el ámbito positivo, se destaca su utilidad en la asistencia tutorizada, acceso constante a la información y evaluación sofisticada. Las repercusiones negativas se destacan especialmente cuando se trata del acceso sistemático a información y servicios disponibles las 24 horas del día. Se advierte que esto puede llevar a aprendizajes no deseados o perjudiciales debido a la falta de intervención educativa humana. Una sección importante del artículo se centra en la evaluación formativa y su supuesta inmunidad al ChatGPT. La autenticidad, la confianza y la definición de la evaluación formativa se exploran detalladamente. Se argumenta que la verdadera evaluación educativa de calidad no puede ser reemplazada por la inteligencia artificial, y se discuten los principios de demostración y confianza en este contexto. Finalmente, el artículo propone una reflexión sobre la necesidad de desarrollar una pedagogía específica para la IA generativa, destacando la importancia de una teoría del aprendizaje adaptada a estas tecnologías emergentes. Se explora el concepto de aprendizaje inteligente en este contexto y se sugiere la necesidad de una adaptación constante en la pedagogía para incorporar eficazmente la IA generativa en los procesos educativos.

*Palabras clave:* Inteligencia artificial generativa, ChatGPT, impacto educativo, evaluación formativa, pedagogía adaptativa.

#### Abstract

This article delves into the impact of generative artificial intelligence (AI), specifically focusing on the ChatGPT model, within the educational sphere. It initiates by exploring the essence of generative programs, such as Transformer AI models, and their implications for education. The fundamental inquiry into what these programs entail and how they utilize statistics from extensive datasets to generate text is thoroughly examined. The article underscores both the positive and negative consequences arising from the integration of ChatGPT in education. On the positive spectrum, its efficacy in tutoring assistance, continual access to information, and advanced assessment capabilities are emphasized. Negative repercussions come into focus, particularly concerning systematic access to information and services available 24/7, cautioning against potential unwanted or detrimental learning outcomes due to the absence of human educational intervention. A significant section of the article centers on formative assessment and its purported resilience against ChatGPT. Authenticity, trust, and the definition of formative assessment are scrutinized in detail, arguing that genuine and high-quality educational assessment cannot be supplanted by artificial intelligence. Principles of demonstration and trust are further discussed in this context. Finally, the article proposes a contemplation on the necessity of formulating a specific pedagogy for generative AI, emphasizing the importance of an adaptive learning theory tailored to these emerging technologies. The concept of intelligent learning is explored in this context, highlighting the need for continuous pedagogical adaptation to effectively incorporate generative AI into educational processes.

*Keywords:* Generative artificial intelligence, ChatGPT, educational impact, formative assessment, adaptive pedagogy.

Recibido: 18/09/2023

Aceptado: 02/11/2023



Este es un artículo de acceso abierto bajo la licencia CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/deed.es>).

## 1. ChatGPT

### Programas generativos y ChatGPT: La aparición del ChatGPT y su impacto

Tras su aparición, los medios de comunicación, los negocios y las finanzas se han hecho eco, con gran alarde de titulares, del impacto del ChatGPT, así Forbes<sup>1</sup> y NBC se reseñan el último informe del Bank of American sobre la importancia de la IA, que compara con la de la electricidad. Y lo hace a partir de un hecho: La aparición de este programa generador y transformador de textos.

Por otra parte, ha llevado a reputados expertos en la sociedad digital, y en la tecnología educativa en particular, a un gran entusiasmo y a la manifestación de expresiones de arrebato como hace tiempo no se habían visto. Así en el campo de la educación digital, Donald Clark ha llegado a decir (2023) que

ChatGPT podría pasar a ser tutor, profesor, formador, instructor, experto, mentor, entrenador sobre cualquier tema imaginable. Tendrán fines educativos específicos; personalizado, experiencia en el tema, apropiado para la edad, sensibilidad cultural y ser infinitamente útil y paciente las 24 horas del día, los 7 días de la semana. Ideas que después ha ampliado en su blog Plan B, en febrero de 2023, diciendo que ChatGPT es un sistema de diálogo y no solo un productor de piezas de texto monolíticas, puede producir cosas maravillosas. Esta es una nueva habilidad para los diseñadores de aprendizaje: la IA generativa permite todo tipo de indicaciones complejas mediante iteraciones, lógica y parámetros que mejoran el resultado. También que permite mejorar el diseño [instruccional]inicial, haciendo funcionar los trabajos y proyectos con sólo las partes interesantes, conseguir objetivos, adquirir competencias y habilidades deseadas, planes de estudio e incluso títulos para sus iniciativas de aprendizaje. Para crear contenido, uno puede solicitar contenido completo con el nivel adecuado de detalle y matiz, resúmenes, imágenes, todo en diferentes estilos y voces, adecuado para diferentes audiencias. Puede crear evaluaciones y tareas con rúbricas completas para corregir, también actividades de aprendizaje para debates, escenarios y juegos de roles.

En esta línea ha elaborado “una lista de 100 ideas rápidas para profesionales del aprendizaje para permitir que estos amplíen su perspectiva sobre esta tecnología.

Por último, asegura que *éste es solo el comienzo de las herramientas de IA que mejorarán drásticamente la productividad de los docentes, los diseñadores de aprendizaje y los alumnos.*

Todo ello sin experimentación alguna y sin evidenciar nada de lo que dice.

Hay otros autores. Entre ellos, en el plano español, es igualmente optimista, Andrés Pedreño<sup>2</sup> que llega a decir como *ideas fundamentales* para la enseñanza universitaria las siguientes:

Como herramientas de un enorme potencial a corto plazo, ChatGPT basado en GPT3 es solo el inicio de una serie de herramientas con un enorme potencial para la docencia y la investigación; En la a docencia, ChatGPT es objeto (sic) de disrupción. Los alumnos llevan en sus bolsillos una herramienta que puede superar al profesor que está en la tarima si éste se limita a dar información convencional sobre los temas de la asignatura. La IA obliga a redefinir la enseñanza. (...) Los exámenes y los sistemas de

---

1 Fuente: Forbes. Suplementó de negocios del 1 de marzo de 2023.

<https://www.blogger.com/blog/post/edit/9065073132855231377/1764959125079006163>

2 <https://www.blogger.com/blog/post/edit/9065073132855231377/1764959125079006163>

evaluación convencional también deberán ser profundamente revisados. La educación personalizada se erige como fundamental para impedir brechas relevantes entre alumnos (incluso entre profesores).

También es una oportunidad y reto. En conclusión, ChatGPT representa una oportunidad para transformar la educación en las universidades y mejorar la experiencia de los estudiantes. Al utilizar tecnología avanzada para procesar y comprender el lenguaje humano, ChatGPT también abre la puerta a nuevos horizontes en la educación, brindando a los estudiantes la oportunidad de aprender de manera más personalizada y adaptativa. A los investigadores les puede ayudar a romper los límites para procesar y asimilar grandes cantidades de información.

Más allá de lo que diremos extensamente en este trabajo, cabe plantearse la gran responsabilidad que supone atribuir en una fase tan temprana unos juicios no suficientemente justificados y temerarios sobre la atribución de todos esos “poderes” y capacidades a ChatGPT.

Con un poco más de base empírica, tras someter a ChatGPT a algunas preguntas que dieron respuestas contradictorias cuando no incorrectas pero que serían resueltas por alumnos de Primaria o de Secundaria que indudablemente echarían mano de recursos elementales de pensamiento abstracto, simbólico, matemático o lógico, llegamos a formularnos algunas preguntas. ¿Cómo, en un contexto real, de clase, con alumnos reales, ChatGPT va a enseñar? ¿Cómo va tutelar, evaluar, ... y hacer todo eso que presuponen esos autores en sus visiones? ¿Cómo va a realizar esas tareas (tutoría, evaluación, diseño instruccional, ...), una aplicación que respondía así a problemas simples?

Obviamente quienes atribuyen a ChatGPT todas estas propiedades y capacidades lo hacen no sólo sin ninguna evidencia en que apoyarse, sino de una forma precipitada. ¿Por qué ese interés en que la IA generativa tenga éxito en presupuestos tan ambiciosos? ¿Qué y quién está detrás de estas situaciones?

En la Figura 1 reproducimos un volcado de pantalla con los resultados de Google Scholar cuando preguntamos por Generative AI education “empirical evidence”.

Figura 1. Consulta en Google Scholar sobre Generative AI education “empirical evidence”

The screenshot shows a Google Scholar search interface. The search bar contains the text "Generative AI education. 'Empirical evidence'". Below the search bar, it indicates "Artículos" and "Aproximadamente 21.000 resultados (0,03 s)". The results are listed in a table-like format with filters on the left. The first result is titled "Are both generative AI and ChatGPT game changers for 21st-Century operations and supply chain excellence?" by SF Wamba, MM Queiroz, CJC Jabbour, et al. The second result is "Why Are Conditional Generative Models Better Than Unconditional Ones?" by F Bao, C Li, J Sun, J Zhu. The third result is "Bayesian generative active deep learning" by T Tran, T Do, I Reid. The fourth result is "Generate the Future of Work through AI: Empirical Evidence from Online Labor Markets" by J Liu, X Xu, Y Li, Y Tan. The fifth result is "Empowering Education through Generative AI: Innovative Instructional Strategies for Tomorrow's Learners".

Artículos	Aproximadamente 21.000 resultados (0,03 s)
<p>Cualquier momento Desde 2023 Desde 2022 Desde 2019 Intervalo específico...</p>	<p>Are both <b>generative AI</b> and ChatGPT game changers for 21st-Century operations and supply chain excellence? SF Wamba, MM Queiroz, CJC Jabbour... - International Journal of ..., 2023 - Elsevier ... of ChatGPT, a <b>Generative Artificial Intelligence (Gen-AI)</b>, has ... However, <b>empirical evidence</b> on <b>Gen-AI's</b> effects in O&amp;SCM ... areas by <b>Gen-AI/ChatGPT, education</b> and healthcare are our ... ☆ Guardar 09 Citar</p>
<p>Ordenar por relevancia Ordenar por fecha</p>	<p>Why Are Conditional <b>Generative</b> Models Better Than Unconditional Ones? [PDF] arxiv.org F Bao, C Li, J Sun, J Zhu - arXiv preprint arXiv:2212.00362, 2022 - arxiv.org ... <b>empirical evidence</b> demonstrates that conditional <b>generative</b> ... that the key of conditional <b>learning</b> is to partition the data ... on why conditional <b>learning</b> of <b>generative</b> models generally ... ☆ Guardar 09 Citar Citado por 3 Artículos relacionados Las 3 versiones 08</p>
<p>Cualquier idioma Buscar solo páginas en español</p>	<p>Bayesian <b>generative</b> active deep <b>learning</b> [PDF] mlr.press T Tran, T Do, I Reid... - ... on Machine Learning, 2019 - proceedings.mlr.press ... with data augmentation – we provide theoretical and <b>empirical evidence</b> (MNIST, CIFAR-10, 100), and SVHN) that our approach has more efficient training and better classification ... ☆ Guardar 09 Citar Citado por 117 Artículos relacionados Las 15 versiones 08</p>
<p>Cualquier tipo Artículos de revisión <input type="checkbox"/> incluir patentes <input checked="" type="checkbox"/> incluir citas Crear alerta</p>	<p>"Generate" the Future of Work through <b>AI: Empirical Evidence</b> from Online Labor Markets [PDF] arxiv.org J Liu, X Xu, Y Li, Y Tan - arXiv preprint arXiv:2308.05201, 2023 - arxiv.org ... This research contributes to the limited empirical repository exploring the profound influence of LLM-based <b>generative AI</b> on the labor market, furnishing invaluable insights for workers, ... ☆ Guardar 09 Citar Citado por 1 Las 5 versiones 08</p>
	<p>Empowering <b>Education</b> through <b>Generative AI</b>: Innovative Instructional Strategies for Tomorrow's Learners [PDF] ijble.com</p>

Google Scholar da resultados por exceso, es decir es generosa atribuyendo el carácter de investigación a TFM, TFG y ensayos evaluativos, pero en este caso el único resultado apreciable es de otro tema, del Teorema de Bayes y *deep learning*.

La consulta está hecha el 14 de septiembre de 2023. La conclusión es: Si bien todo lo que se dice sobre IA generativa y ChatGPT en relación con educación y aprendizaje merece respeto, demostrado lo que se dice demostrado, no hay nada.

### **¿Qué son los programas generativos “Transformer” AI, o los generadores de texto a partir de estadísticas de grandes cantidades de datos?**

Para lo que nos ocupa conviene destacar en primer lugar que cuando se habla de ChatGPT se habla de dos cosas. Aunque de forma banal, en nuestro entorno casi todos hablan de la primera e ignoran, se supone que no deliberadamente, la segunda. Lo cual es grave, sobre todo cuando se habla desde una perspectiva educativa.

En primer lugar, se trata de un modelo de predicción de texto transformador y generativo. Es una enciclopedia. Pero también **es un modelo conversacional basado en un modelo instruccional**. Es decir, es un modelo que se basa en la conversación para su aprendizaje propio, mediante el cual incorpora criterios. Esto es lo que todo el mundo desconoce u olvida.

En la primera acepción, GPT son las siglas de *Generative Pre-trained Transformer*. Es un modelo de predicción de texto gigante. Está entrenado por OpenAI en el uso de 500 mil millones de palabras. Su precedente, el GPT-3, y él no solo son capaces de escribir correctamente en varios idiomas, sino que sobre todo es un modelo enciclopédico, en esta acepción, que integra una gran cantidad de referencias y datos del mundo real (personas, eventos, conocimiento científico). Hasta ahora, su precedente GPT-3 nunca se abrió al público en general por razones de riesgo. GPT-3 no tenía un mecanismo de inhibición real basado en experiencia y valores humanos y podía generar cualquier disparate, siempre que el texto sea solo superficialmente coherente.

Según la segunda acepción, ChatGPT se basa en InstructGPT, un modelo conversacional de aprendizaje. Lo importante, para nuestro propósito y nuestro enfoque del tema, es que está basado en el aprendizaje por refuerzo a partir de la retroalimentación humana (RLHF). Es una *affordance* creada a partir de la anotación de textos generados. Pero lo importante es que el modelo incorpora toda una serie de recompensas y penalizaciones con la función de reforzar la coherencia del texto generado, evitar falsedades flagrantes, pero también moderar por anticipación de posibles derivas tóxicas. Mediante estas funciones, ChatGPT, de entre todo inmenso repertorio de de respuestas lingüísticamente correctas que se podrían generar, ChatGPT opta por aquellas que son más correctas en un contexto de chat: esto se llama “alineación de IA”. Basado en este mismo principio, ChatGPT a veces también se niega rotundamente a responder.

Esta doble naturaleza condiciona y hace comprender la esencia y las limitaciones de ChatGPT. La distinción entre modelo textual y modelo conversacional es fundamental para entender cómo funciona:

- El **modelo lingüístico no está actualizado** y no se alimenta de las sugerencias de los usuarios. La formación, la actualización de las bases de datos es muy cara y solo se renovará una vez al año por término. Esta es la razón por la que hasta muy recientemente (Langlais, 7 febrero 2023) no hacía referencias al mundo después de 2021. No obstante, hemos comprobado que algunas aportaciones hechas en discusiones recientes han sido incorporadas, pero ha sido más en temas de discusiones procedimentales que en nuevos datos (enfoque conversacional).

- El **modelo conversacional** continúa refinándose en función de las contribuciones de los usuarios cada 3 o 4 semanas en promedio. Con ellas OpenAI mejora la alineación del modelo con la intención del usuario o la plataforma. Ante la paralización del modelo textual, el modelo conversacional hace generaciones de mayor calidad y, como consecuencia, de su metodología de refuerzo, penaliza más a las generaciones no deseadas. La última versión, fechada el 30 de enero último, penaliza más “alucinaciones”. Término utilizado para designar generaciones puramente poco o nada reales o posibles).

### ¿Qué son?

Los programas transformadores en IA son una modalidad de arquitectura de redes neuronales que ha ido ganando popularidad. A su vez, las redes neuronales son redes de nodos que simulan las redes de neuronas biológicas (no necesariamente humanas). Y tienen la virtud de aprender de una determinada forma: Las redes neuronales aprenden (o son entrenadas) procesando ejemplos, cada uno de los cuales contiene una "entrada" y un "resultado" computables como sucesos probabilísticos, de manera que se pueden establecer asociaciones ponderadas de probabilidad entre los dos (en esto tiene mucha importancia el análisis bayesiano), que se almacenan dentro de la estructura de datos de la propia red.

Los transformadores han sido puestos de moda por OpenAI en los modelos de lenguaje precedentes de ChatGPT, la serie GPT, y en el propio ChatGPT, pero también fueron utilizados antes sin tanta resonancia, como por ejemplo por DeepMind para AlphaStar, su programa que derrotó a un jugador profesional de Starcraft.

En general los programas transformadores IA se desarrollaron para resolver el problema conocido como de la *transducción* de secuencias (Graves, 2012) (transduction sequence en inglés) o traducción automática neuronal. Eso tiene que ver con cualquier tarea automática de dispositivo digital que transforme una secuencia de entrada en una secuencia de salida de carácter distinto. Esto incluye reconocimiento de voz, transformación de texto a voz, reconocimiento de imágenes, etc.

El transformador IA GPT-3, y después ChatGPT, no suponen ninguna novedad conceptual más allá de la potencia de sus bases de datos y la capacidad de sus algoritmos para procesar esa gran cantidad de datos rápidamente. De ahí lo de grande en *large language models* (LLM) como de otra forma se llaman. Funcionan como un *completador* de texto altamente capacitado, similar al que se encuentra en los teléfonos móviles y las interfaces de correo electrónico. Lo que sucede es que en lugar de mirar hacia atrás a los últimos caracteres y usarlos para predecir la próxima palabra o dos, presta atención a las 700 u 800 palabras anteriores que ha escrito, para continuar escribiendo una historia corta completa, un blog o un ensayo de estudiante. Ese mismo tipo de programa también puede resumir un artículo científico en un lenguaje más simple, escribir una reseña, traducir idiomas y responder preguntas generales. En resumen, un *Transformer AI* es una herramienta de lenguaje de propósito general.

Y éste va a ser su gran déficit para lo que estamos tratando, la educación.

“Los grandes modelos de lenguaje preentrenados no se basan en otros dominios de la experiencia, como el video o la interacción física del mundo real, y por lo tanto carecen de una gran cantidad de contexto sobre el mundo” (Brown et al., 2020, p. 34). Los transformadores son modelos de lenguaje, no de conocimiento a través de la experiencia, la asignación de sentido o los valores. No están diseñados para ser humanos reflexivos, ecuanimes y éticos, o en particular para ser académicos, para verificar referencias académicas y garantizar que la evidencia esté basada en hechos. En términos humanos, son esencialmente inexpertos, irreflexivos y amorales. No tienen la capacidad de reflexionar sobre lo que han escrito, ni de juzgar si es exacto y decente.

## **Conclusiones generales sobre ChapGPTG y otros Large language models**

Coincidiendo con Sharples (2022), pensamos que las características enunciadas en el apartado anterior harán que ChatGPT tenga poco futuro en la Educación. No podemos poner en las manos de un agente deseablemente educativo, sea chatbot u otro, la posibilidad de que suministre a los alumnos un contenido hipotéticamente inapropiado o del que no podemos garantizar que no lo sea. En todo caso lo podríamos poner en manos de profesionales consolidados con un juicio, capacidad profesional y experiencia que le permita discernir y evaluar los contenidos ofrecidos por ChatGTP más allá de lo que ofrece como resultado de una consulta.

Consciente de ello, OpenAI ha proporcionado un complemento para GPT-3 que filtra el lenguaje inapropiado. Sin embargo, es poco probable por el costo que supone en relación con la rentabilidad que la empresa produzca herramientas para verificar la precisión de los resultados en relación en todo caso con la exactitud de las fuentes de donde se toman los datos o las afirmaciones evaluando su rigor como lo podría hacer un humano. De esta manera, como hemos dicho, el enfoque y la aplicación estaría en la inteligencia artificial general apoyando trabajos profesionales, no en la educación.

Quizás otras empresas, en un futuro, podrían suministrar herramientas para verificar la precisión y la fiabilidad de las referencias generadas o para agregar referencias específicas para un artículo. Pero en todo caso estas herramientas no superarían la limitación fundamental de los modelos de los lenguajes Transformer, como ChatGPT: que no tienen un modelo interno evaluable y accesible para modificar en función de cómo funciona el mundo, proporcionando una base de conocimientos para que el sistema reflexione sobre la precisión y la sapiencia de su trabajo generado.

Gunning et al (2019) están realizando investigaciones para desarrollar una IA neuronal explicable y Garcez y Lamb (2020) están investigando sistemas híbridos de IA neuronal/simbólica que podrían abordar este problema.

En todo caso coincidimos con Sharples (2023) cuando en un tuit dice:

He visto a algunos académicos desconcertados sobre por qué ChatGPT inventa referencias académicas y por qué su "base de datos" se detiene en 2021. Ese es un malentendido fundamental de cómo funciona. Es un modelo de lenguaje, no una base de conocimiento. Produce "continuaciones" no "hechos". Necesitamos una nueva alfabetización en IA.

Como ideas generales pensamos que además de lo dicho esta línea de desarrollo, para ser realmente eficiente y para su puesta en práctica de forma generalizada, más allá de las intenciones enunciadas por los posts y notas de prensa de sus departamentos de comunicación y marketing, debe profundizar en

A) Utilizar como herramientas más finas las que se derivan del nuevo pensamiento bayesiano que expusimos (Zapata-Ros, M. & Palacios, Y. B., 2021), para seleccionar en función de una analítica de datos, para la cual tienen ya potentes bases, las respuestas en los informes a las preguntas que se les hacen.

B) Tener presente elementos más finos de lo que son procesos cognitivos de aprender y sobre qué es aprender.

Respecto de lo primero, el análisis y razonamiento probabilístico es un subcampo del aprendizaje automático aplicado al procesamiento del lenguaje natural (PNL). Y, en su contexto, un campo de Probabilidad, la estadística bayesiana, puede ofrecer técnicas únicas para el PNL.

Como en el resto de la tradición bayesiana, pero ahora apoyada por el análisis de grandes conjuntos de datos, la asignación de probabilidad a un suceso se basa en la probabilidad de su inverso (probabilidad a priori), a través del resultado en experimentos conocidos (probabilidad inversa, probabilidad compuesta y teorema de Bayes). En contraposición a la otra gran escuela clásica, la asignación a través de la frecuencia (probabilidad determinista o frecuentista) (Drury, agosto 2019):

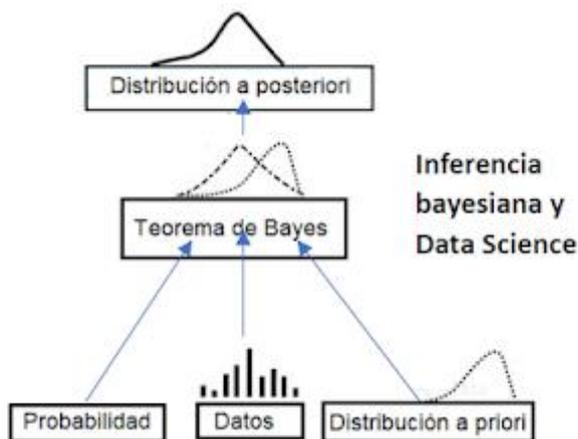
Cohen (2019) afirma que el objetivo de la estimación puntual bayesiana es: "resumir la parte posterior de los parámetros en un conjunto fijo de parámetros", y vincula este objetivo a un enfoque frecuentista conocido como estimación de máxima verosimilitud.

En otro trabajo Barrow (2019), glosando el libro de Cohen (2019) dice:

El procesamiento del lenguaje natural (PNL) experimentó una profunda transformación a mediados de la década de 1980 cuando cambió para hacer un uso intensivo de corpora y de técnicas basadas en datos orientadas al análisis del lenguaje. Desde entonces, el uso de técnicas estadísticas en PNL ha evolucionado de varias formas. Un ejemplo de evolución tuvo lugar a finales de la década de 1990 o principios de la de 2000, cuando se introdujeron las herramientas bayesianas, en toda regla, en el PNL. Este enfoque bayesiano del PNL ha llegado a suplir resolviendo varias deficiencias en el enfoque frecuentista enriqueciéndolo, especialmente en el entorno no supervisado, donde el aprendizaje estadístico se realiza sin ejemplos de predicción de objetivos.

En esquema, este proceso de asignación bayesiana de la probabilidad apoyada con analítica de grandes datos sería:

Figura 2. Proceso de asignación bayesiana de la probabilidad



Fuente: Imagen obtenida de Zapata-Ros, M. & Palacios, Y. B. (2021)

Hemos visto este uso de la probabilidad bayesiana en la pandemia, en modelos de epidemiología, para determinar modelos de evolución de datos sobre contagio y fallecimientos en el COVID y en el procesamiento del lenguaje natural. Pero igualmente podríamos verlo en multitud de casos en los más variados campos científicos y de análisis de procesos. De esta forma, con la automatización de los métodos bayesianos y el uso de modelos gráficos probabilísticos es posible identificar patrones y anomalías en voluminosos conjuntos de datos en campos tan diversos como son los corpus lingüísticos, los mapas astronómicos, añadir

funcionalidades a la práctica de la resonancia magnética, o a los hábitos de compra con tarjeta, online o smartphones. Esto por señalar sólo algunos casos, como son los que se asocian con el análisis de grandes datos y la teoría bayesiana. Pero hay muchos más (Bundy, 2007 y Wing, 2008, a través de Zapata-Ros, 2020 agosto).

Este esquema está presente en multitud de casos. A lo dicho en anteriores ocasiones hay que añadir ahora la idea de pensamiento computacional generalizado (pervasive computational thinking) como avance y desarrollo de lo tratado en el pensamiento computacional simple, el que Wing (2006) plantea en su primera aproximación.

Abundan evidencias sobre la influencia del pensamiento computacional apoyado por el análisis bayesiano en otros campos: Está transformando las estadísticas, donde con el aprendizaje automático, la automatización de los métodos bayesianos y el uso de modelos gráficos probabilísticos es posible identificar patrones y anomalías en voluminosos conjuntos de datos. Esto sería deseable que se incorporase también en futuros Large language models y otros modelos de transformadores de lenguaje

Respecto de lo segundo, y a propósito del ejemplo del pez y de su cabeza, convendría tener presente y como base de estudios y desarrollos un tratamiento más riguroso y detallado de los avances de la psicología cognitiva, y no quedarse en teorías obsoletas como son las del aprendizaje por refuerzo, y de las teorías instruccionales sobre lo que son procesos cognitivos de aprender y sobre qué es aprender y cómo se utilizan en el diseño educativo. Lo hemos visto antes y también lo veremos más adelante, cuando hablemos de condiciones del aprendizaje, comprender, atribuir sentido, incorporación, autonomía y transferencia. Y de cómo ello debe ser garantizado por una autentica evaluación, de manera formativa y con dominio en el aprendizaje.

## **Repercusiones para la educación y el aprendizaje**

### **Repercusiones positivas**

Bastantes cosas de las que se han dicho sobre las ventajas del ChatGPT en educación, matizadas y contextualizadas adecuadamente, son ciertas.

Así por ejemplo es obvio que este *transformer* como el resto forman parte de la IA de asistencia educativa tutorizada. De cuya naturaleza educativos participan. También es previsible que compartan esa naturaleza las versiones sucesivas ChatGPT-n. Lo que se dice de esta AI es cierto para ellos: proporcionarían acceso de forma sistematizada a la información existente y a autoservicios de ese tipo, de forma permanente las 24 horas del día, los 7 días de la semana. Todo ello si no entramos más en detalle de como lo hacen, con qué fiabilidad, etc tal como lo hemos visto y lo seguiremos viendo.

También serán útiles para cierto tipo de evaluación. Sobre todo, si consideramos que unas pruebas “objetivas” del tipo de respuesta múltiple o de verdadero y falso, más sofisticadas que las actuales en el sentido de tener más en cuenta condiciones de contexto, se pueden considerar evaluación de aprendizajes y de logros.

Así, a lo largo de muchos años, la IA que apoya a los estudiantes se ha desarrollado para incluir, por ejemplo, herramientas de aprendizaje adaptativo “para dominios complejos como lenguajes de programación, matemáticas, medicina, física, solución de problemas de aviónica y electrónica” (Wasson 1997:572); también para la captura y el análisis de una amplia gama de señales en el aula (por ejemplo, medir la atención, la empatía y la emoción), el uso de una gama cada vez mayor de hardware (desde teléfonos móviles hasta auriculares EEG), en cuyo contexto y con cuya ayuda se han diseñado chatbots para brindar a los alumnos soporte todo el tiempo, y por último , para incluir organizadores de redes de aprendizaje diseñados para construir comunidades de estudiantes, evaluación automática de escritura, etc.

En general en las últimas tres décadas, la mayor parte del enfoque de investigación sobre IA y Educación (AIED) se ha centrado en el apoyo que la IA puede suministrar al del alumno, pensando en aquella que, por definición, tiene como objetivo automatizar las funciones del docente, de modo que los alumnos puedan aprender independientemente que tengan profesor o no. Es decir, que tengan su propio tutor personal artificial y puedan aprovechar en sentido muy relajado lo que se sostiene acerca de la enseñanza mentorizada a partir de las conclusiones del estudio conocido como problema de dos sigmas de Bloom.

Sin embargo, gran parte de todo esto adopta un enfoque bastante tosco y primitivo de la pedagogía y, con demasiada frecuencia, se centra en la automatización de prácticas pedagógicas poco eficientes para satisfacer a usuarios poco exigentes, cuando no banales, en lugar de la esforzarse por una innovación de pedagogía científica de calidad, como es por ejemplo de idear formas innovadoras de evaluar y acreditar el aprendizaje, en lugar de facilitar exámenes sin abandonar la idea del propio examen como verificación de registro memorístico o de información, pero cada vez con recursos de verificación de esa información, en vez de considerar la verificación como evaluación del aprendizaje real.

Un beneficio claro está siendo la emergencia y la constatación de la necesidad de otras ideas positivas alternativas para utilizar los *transformers* o con ellos. Así podemos ver que en los próximos tiempos podrían ocurrir algunas de estas cosas:

El ensayo, trabajo fin de máster (TFM), de Grado (TFG), de asignatura, o como quiera que se le llame o se utilice, perdería vigencia. Los profesores podrían restringir las asignaciones por estos conceptos, en las ponderaciones que se hagan para superar estudios, a los trabajos informes o de ensayos, en favor de otro tipo de pruebas como son exámenes supervisados, o trabajos y prácticas muy supervisadas con entrevistas. El problema es que estos procedimientos no generan pruebas, son poco formales, y consumen mucho tiempo.

Otra alternativa que podría generarse es la de establecer tareas escritas reflexivas y contextualizadas que la IA no podría generar. Para esto es clave la interacción. De manera que los sucesivos refinamientos tengan que hacer alusión a las propuestas de mejora específicas de profesor. Por ejemplo, el tutor podría asignarle a cada estudiante un proyecto de investigación independiente, luego solicitar un informe escrito sobre ese proyecto concreto, dar retroalimentación al estudiante sobre el informe y luego pedirle que escriba una reflexión crítica sobre la retroalimentación y los problemas planteados por el proyecto a los que aludió en su primer informe de revisión.

En fin, existen muchas formas imaginativas y nuevas, algunas de las cuales pueden integrar incluso los programas Transformers. La cuestión es que en todas estas como en las anteriores se exigen un considerable esfuerzo por los profesores y un pacto tácito de confianza y de colaboración por parte de alumnos y familias.

Así, una forma imaginativa de incorporar texto generado por estos procedimientos en las tareas educativas podría ser que el profesor emplee un *Transformer AI* para generar un conjunto de ensayos alternativos sobre un tema, luego pida a los estudiantes que los critiquen y escriban sus propias versiones mejores.

O que establezca una pregunta compleja y luego pida a cada estudiante que genere respuestas de IA a la pregunta y que sea el estudiante el que evalúe estas respuestas en relación con los criterios de evaluación.

Seguro que en la literatura y en las prácticas experimentales aparecerán opciones de uso educativo de los *Transformers*, pero nunca tendrán los efectos taumatúrgicos que ahora se les otorgan.

## **Repercusiones negativas**

Se puede decir que tiene tantas repercusiones negativas como positivas hemos visto, siempre que no se no se satisfagan unas mínimas condiciones de intervención educativa humana.

Así cuando hablamos de acceso de forma sistematizada a información existente o de proporcionar autoservicios que funcionan las 24 horas del día, los 7 días de la semana las consecuencias pueden no ser educativas, sino que se pueden incluso derivar en aprendizajes no deseados o perjudiciales, a consecuencia de ausencia de metacognición o de criterios adecuados de administrar la información y de valorarla críticamente. Este efecto también puede producirse a consecuencia de una ausencia de elaboración propia del alumno, con la falta de una adecuada atribución de sentido y, la ya señalada, de interacción que valide la corrección de las respuestas. Cuestión que, como hemos visto, en sentido educativo, ChatGPT no tiene ni se espera que lo tenga próximamente, por no estar en la agenda y ser muy caro. Lo primero como consecuencia de lo segundo.

## **Cuestiones en las que ChatGPT participa de las características de la IA educativa**

Por último, no conviene olvidar que los Transformers y el ChatGPT participan de todos aquellos factores que hacen criticables la IA por tomar decisiones basadas en resultados algorítmicos. Es decir, de todos los inconvenientes que en general tienen que ver con la IA como generadora de ideas sin mediación ni validación humana intermedia.

## **En educación, esto es aplicable igualmente a ChatGPT**

Surden (2019) identifica problemas de la sociedad contemporánea que incluyen el sesgo en la toma de decisiones algorítmicas, en la interpretabilidad de sistemas de muy variada naturaleza a través de la IA. Y que son derivados de la falta de transparencia sobre cómo los sistemas de IA toman sus decisiones y en la deferencia que se hace hacia ellos como con una especie de respeto y de reverencia que los hace como no cuestionables. Todos ellos considerados como problemas en línea con los desafíos reconocidos para la IA y su aplicación en general.

Este razonamiento y estas conclusiones las hace Surden cuando describe la aplicación de la IA en el uso y estudio de las leyes. Pero cada uno de estos temas contemporáneos es, de forma análoga, extremadamente relevante para comprender y evaluar el uso de la IA en la educación, en particular a la luz de lo que pueda surgir con ChatGPT.

*No olvidemos que en su origen es un producto de la psicología conductista más pura, del aprendizaje por refuerzo*

Cuando creíamos que los modelos simplistas y mecanicistas de la psicología conductista y del aprendizaje humano por refuerzo ya se habían extinguido y se consideraban como una cosa del pasado, ahora resucitan con los Transformers GPT.

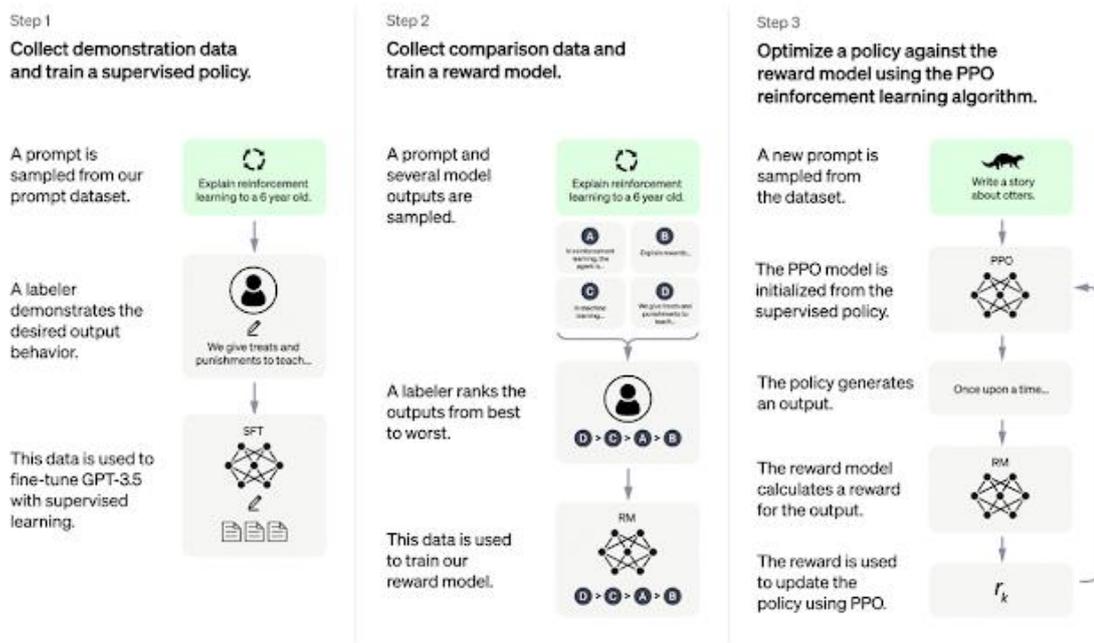
Recordemos que, según toda la literatura suministrada por Open AI, ChatGPT se basa en InstructGPT, un *modelo conversacional* de aprendizaje por refuerzo a partir de la retroalimentación humana (*RLHF*). Es una versión “rectificada” de GPT-3, creada a partir de procesos de anotación genérica de textos, que genera otros textos. De esta forma, el modelo incorpora para ello toda una serie de recompensas y penalizaciones que cumplen varias funciones: reforzar la coherencia del texto generado, evitar falsedades flagrantes, pero también moderar por anticipación posibles derivas tóxicas.

Así pues, el avance de ChatGPT sobre los GPT3 y anteriores se basa en que, de todo el universo de respuestas lingüísticamente correctas que podrían generar los predecesores,

chatGPT opta por aquellas que son más correctas en un contexto de chat: esto se llama “alineación de IA”. Utilizando este mismo principio, chatGPT a veces también se niega rotundamente a responder.

Éste es el esquema:

Figura 3. Pasos para alinear modelos lingüísticos para seguir instrucciones



Fuente: Ryan Lowe & Jan Leike (January 27, 2022). Aligning language models to follow instructions. Language, Human feedback, Safety & Alignment, Responsible AI, Milestone, Publication. <https://openai.com/research/instruction-following>

Este trabajo (Ryan Lowe & Jan Leike, January 27, 2022) enlaza con otro documento del propio Open IA<sup>3</sup>, *Aligning language models to follow instructions*, que hace referencia al aprendizaje por refuerzo: InstructGPT.

En este documento y en otros se señala que, para enseñar a los tres modelos (Chap GPT, GPT 3 e InstructGPT), se utilizaron formas semejantes a las conductistas o de aprendizaje por condicionamiento o por refuerzo para enseñar a los programas, para que aprendiesen por estos métodos desfasados.

Dando por supuesto eso, nos podemos plantear ¿se puede enseñar sin saber aprender? ¿Podrían estos programas actuar como educadores si están hechos sin conocer los modelos y teorías vigentes y probadas del aprendizaje? ¿Qué garantías ofrecen? Eso cuestionaría todo el uso educativo o al menos como tutor o instructor o evaluador de ChatGPT.

De hecho, el aprendizaje por refuerzo hace que el ChatGPT esté en una fase histórica del desarrollo del aprendizaje previa a las ideas y a los planteamientos cognitivos.

Al programa se le enseña por refuerzo, pero carece de habilidades, métodos o formas de operar (de hecho, los programadores ni se lo plantean en sus algoritmos, no hay una pedagogía y mucho menos una pedagogía cognitivista) que tengan en cuenta:

- *La atribución de sentido*
- *La experiencia o experiencias vividas e incorporadas.*
- *Los valores.*

3 <https://openai.com/research/instruction-following>

Tampoco cuestiones, desarrollos y condiciones que implican considerar el conocimiento estructurado por relaciones de significación necesarias para ensamblar, integrar y cohesionar los nuevos conocimientos con los previos. Es lo que se llama el andamiaje cognitivo, los conceptos inclusores, los conceptos previos, las relaciones de significación, etc.

En definitiva, a ChatGPT le falta el alma o lo que Siemens consideraba el “beingness”

Este procedimiento, el de recurrir a la imitación humana del aprendizaje por refuerzo, lo podemos ver en las propias declaraciones de OpenIA:

En *Introducing ChatGPT* dice, en el apartado “métodos”:

Entrenamos este modelo usando Aprendizaje por refuerzo a partir de retroalimentación humana (RLHF), usando los mismos métodos que InstructGPT, pero con ligeras diferencias en la configuración de recopilación de datos. Entrenamos un modelo inicial mediante un ajuste fino supervisado: los entrenadores humanos de IA proporcionaron conversaciones en las que jugaron en ambos lados: el usuario y un asistente de IA. Les dimos a los capacitadores acceso a sugerencias escritas en modelos para ayudarlos a redactar sus respuestas. Mezclamos este nuevo conjunto de datos de diálogo con el conjunto de datos InstructGPT, que transformamos en un formato de diálogo.

Para crear un modelo de recompensa para el aprendizaje por refuerzo, necesitábamos recopilar datos de comparación, que consistían en dos o más respuestas del modelo clasificadas por calidad. Para recopilar estos datos, tomamos conversaciones que los entrenadores de IA tuvieron con el chatbot. Seleccionamos al azar un mensaje escrito por un modelo, probamos varias finalizaciones alternativas e hicimos que los entrenadores de IA las clasificaran. Usando estos modelos de recompensa, podemos ajustar el modelo usando la Optimización de Política Proximal. Realizamos varias iteraciones de este proceso.

En la descripción que se hace de InstructGPT en el documento *Aligning language models to follow instructions* de OpenIA, dice:

Para hacer que nuestros modelos sean más seguros, más útiles y más alineados, utilizamos una técnica existente llamada aprendizaje por refuerzo a partir de la retroalimentación humana (RLHF). En las solicitudes enviadas por nuestros clientes a la API, nuestros etiquetadores brindan demostraciones del comportamiento deseado del modelo y clasifican varios resultados de nuestros modelos. Luego usamos estos datos para ajustar GPT-3.

Los modelos InstructGPT resultantes son mucho mejores para seguir instrucciones que GPT-3. También inventan hechos con menos frecuencia y muestran pequeñas disminuciones en la generación de productos tóxicos. Nuestras etiquetadoras prefieren las salidas de nuestro modelo 1.3B InstructGPT a las salidas de un modelo 175B GPT-3, a pesar de tener más de 100 veces menos parámetros. Al mismo tiempo, demostramos que no tenemos que comprometer las capacidades de GPT-3, según lo medido por el rendimiento de nuestro modelo en las evaluaciones académicas de PNL.”

Y también, en ese mismo documento, sobre InstructGPT () dice

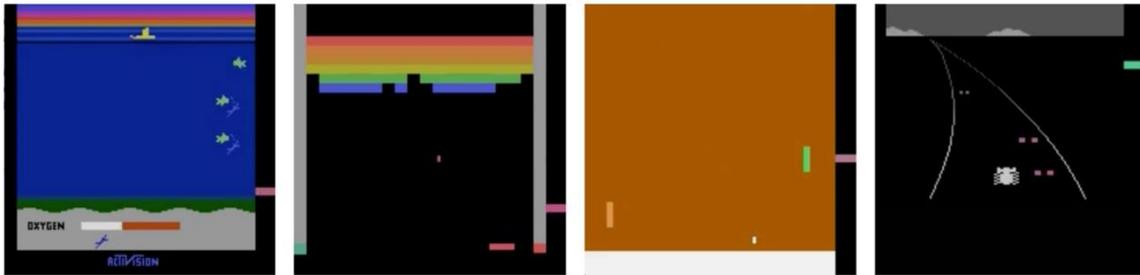
Para entrenar modelos de InstructGPT, nuestra técnica central es el aprendizaje por refuerzo a partir de la retroalimentación humana (RLHF), un método que ayudamos a ser pionero en nuestra investigación de alineación anterior. Esta técnica utiliza las preferencias humanas como una señal de recompensa para ajustar nuestros modelos, lo

cual es importante ya que los problemas de seguridad y alineación que buscamos resolver son complejos y subjetivos, y no se capturan completamente mediante métricas automáticas simples.

Por último, en el documento “Learning from human preferences“, dan una vuelta de tuerca más al decir que se basan en los mecanismos simples, que utilizan los humanos para discernir entre dos procedimientos, que utilizan los juegos de Atari, en un alarde de conductismo primario:

Hemos probado nuestro método en una serie de tareas en los dominios de robótica simulada y Atari (sin tener acceso a la función de recompensa: en Atari, sin tener acceso a la puntuación del juego). Nuestros agentes pueden aprender de los comentarios humanos para lograr un rendimiento sólido y, a veces, sobrehumano en muchos de los entornos que probamos. En la siguiente animación puedes ver agentes entrenados con nuestra técnica jugando a una variedad de juegos de Atari.

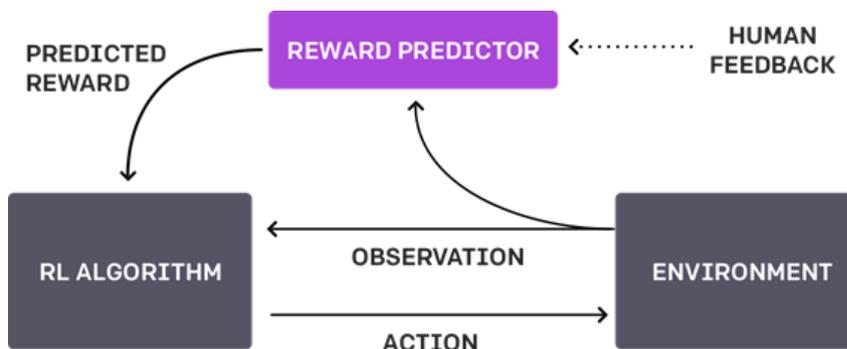
Para a continuación ilustrar el tema con animaciones de juegos elementales del tipo:



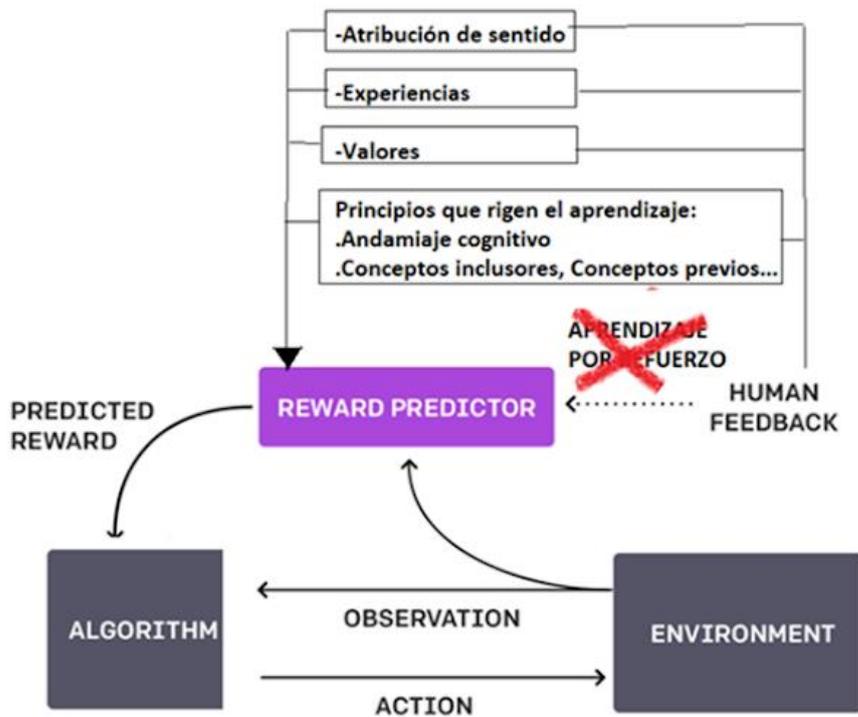
Imagino que tanto Clark como Pedreño habrán visto esto antes de realizar las declaraciones que señalábamos al principio de este documento (serie de posts).

No todo es descartable, en “aprender de las preferencias humana” hay un esquema interesante al menos la parte suficiente para justificar que es compendio de métodos e ideas conductistas los que se utilizan para enseñar a ChatGPT. Es un ciclo de retroalimentación de 3 pasos entre el ser humano, la comprensión del objetivo por parte del agente y la capacitación por RL (aprendizaje por refuerzo restringido al mundo de los videojuegos, Reinforcement Learning RL) (Figura 4).

Figura 4. Ciclo de retroalimentación del ChatGPT



Que nos sugiere este otro:



### Ensayos. TFM, TFG, tesis doctorales y artículos científicos. La evaluación con informes, ensayos y proyectos

Los ensayos, trabajos o informes han formado parte de la evaluación educativa, y principalmente académica y universitaria, desde principios del siglo XIX, cuando los estudiantes de algunas universidades europeas debían escribir artículos académicos para su discusión en seminarios (Kruse, 2006).

Hoy los ensayos académicos continúan siendo un pilar de la evaluación en escuelas, colegios y universidades muy aceptables porque superan las limitaciones de los exámenes, restringidos a lo que un alumno sabe o es capaz de recordar en un momento determinado. Han triunfado y constituyen los conocidos TFM, TFG y trabajo final de cualquier curso o asignatura. Incluso una disertación de tesis doctoral no escapa a este formato. Son fáciles de establecer, permiten evaluar la profundidad de la comprensión e hipotéticamente demuestran la capacidad de los estudiantes sobre cómo expresar un argumento. En los últimos años, el alcance de los trabajos escritos se ha ampliado para incluir trabajos narrativos, argumentativos, reflexivos, expresivos, receptivos y analíticos.

Sin embargo, no todo es bueno, la práctica generalizada de establecer tareas escritas como trabajo de curso ha sido criticada como laboriosa e injusta (Race, 2018). Pero sobre todo está sujeto a trampas, imposturas y fraudes en general. Prácticas que actualmente se han constituido en un lucrativo negocio con contratos de encargo a través de "fábricas de TFM, TFG y trabajos en general", que venden tareas de este tipo escritas por encargo, a tarifas de hasta 500 euros por escribir un ensayo de 20 páginas. Un estudio de Newton (2018) encontró que el 15,7% de los estudiantes encuestados admitió haber pagado a otra persona para que escribiera una tarea.

Ahora la cosa con ChatGPT da un salto cualitativo: Una consecuencia, no deseada ni perseguida como objetivo, de estos sistemas es que democratizan el fraude. Un estudiante puede generar un ensayo completo en segundos, a un costo de alrededor de 50 centavos de dólar.

Peor aún, si cabe, es que los aspirantes a investigadores científicos puedan recurrir a estos sistemas para generar artículos y enviarlos a congresos y revistas.

Desde el lanzamiento de ChatGPT, los investigadores han estado lidiando con los problemas éticos que rodean su uso, porque según Holly Else en *Nature*, “gran parte de su producción puede ser difícil de distinguir del texto escrito por humanos”.

Se pueden citar distintos casos y experimentos:

Uno es el de los investigadores Blanco-González et al. (2022) y O'Connor y ChatGPT (2023) que han publicado un preprint y un editorial escrito por ChatGPT.

Otro es un grupo dirigido por Catherine Gao en la Universidad Northwestern en Chicago, que ha utilizado ChatGPT para generar resúmenes de trabajos de investigación artificiales para probar si los científicos pueden detectarlos. Son interesantes los planteamientos y resultados.

Los investigadores le pidieron a ChatGPT que escribiera 50 resúmenes de investigación médica basados en una selección publicada en *JAMA*, *The New England Journal of Medicine*, *The BMJ*, *The Lancet* y *Nature Medicine*. Luego los compararon con los resúmenes originales pasándolos por un detector de plagio y un detector de salida de IA, también le pidieron a un grupo de investigadores médicos que detectaran los resúmenes artificiales.

El resultado que dio sobre los resúmenes generados por ChatGPT el verificador de plagio sobre la originalidad fue del 100 %, lo que indica que no se detectó plagio. El detector de salida de IA detectó el 66 % de los resúmenes generados. Pero los revisores humanos no lo hicieron mucho mejor: identificaron correctamente sólo el 68 % de los resúmenes generados y el 86% de los resúmenes genuinos. Identificaron incorrectamente el 32 % de los resúmenes generados como reales y el 14 % de los resúmenes genuinos como generados. Todo ello muy elocuente.

Las conclusiones de los autores del preprint, Gao y sus colegas, son

- “ChatGPT escribe resúmenes científicos creíbles”
- “Quedan por determinarse los límites del uso ético y aceptable de grandes modelos lingüísticos para ayudar a la redacción científica”.

Otros trabajos sobre el tema los podemos encontrar en (Stokel-Walker, 2022) *AI bot ChatGPT writes smart essays — should professors worry?* y en (Grove, 2023) *The ChatGPT revolution of academic research has begun*.

En enero de 2023, *Nature* informó sobre dos preprints y dos artículos en los campos de la ciencia y la salud que incluyeron a ChatGPT como autor firmado. Cada uno de estos incluye una afiliación para ChatGPT, y uno de los artículos incluye una dirección de correo electrónico para el "autor" no humano. Según *Nature*, la inclusión de ChatGPT en ese artículo en la firma del autor fue un "error que pronto se corregirá" (Stokel-Walker, 2023). Sin embargo, estos artículos y sus “autores” no humanos ya han sido indexados en PubMed y Google Scholar.

Experiencias como ésta han llevado a las instituciones, revistas y editoriales a tomar medidas. De entre las más importantes podemos destacar lo que sigue.

*Nature* ha definido una política para guiar el uso de modelos de transformadores de lenguaje a gran escala en la publicación científica: prohíbe nombrar herramientas como "autor acreditado en un artículo de investigación" porque "la atribución de autoría conlleva responsabilidad por el trabajo, y Las herramientas de IA no pueden asumir tal responsabilidad”.

La política editorial también aconseja a los investigadores que utilizan estas herramientas que documenten este uso en las secciones Métodos o Agradecimientos de los manuscritos.

Otras revistas (Science y Taylor & Francis) y organizaciones (WAME) están desarrollando urgentemente políticas que prohíben la inclusión de estas tecnologías no humanas como "autores". Van desde prohibir la inclusión de texto generado por ChatGPT o similares en los originales, a exigir total transparencia, responsabilidad y rendición de cuentas sobre cómo se utilizan y se informan dichas herramientas en publicaciones académicas.

La Conferencia Internacional sobre Aprendizaje Automático (ICML *Fortieth International Conference on Machine Learning*) también ha anunciado una nueva política: "Los artículos que incluyen texto generado a partir de un modelo de lenguaje a gran escala (LLM) como ChatGPT están prohibidos a menos que el texto producido se presente como parte del análisis experimental del artículo".

Y así sucesivamente.

Un ejemplo de normas concretas nos lo presenta JAMA, sobre responsabilidades de autores, materiales a publicar, etc. (Instructions for Authors. *JAMA*. Updated January 30, 2023.)

En conclusión, con todo lo anterior, será difícil ignorar el creciente número de estudiantes e investigadores que envían tareas escritas por ChatGPT o los programas similares que vengan.

En el caso de estudiantes, ¿podrá hacerlo Turnitin u otros productos similares? Aquél ha admitido que "ya estamos viendo los comienzos de la próxima ola de IA... cuando los estudiantes pueden presionar un botón y la computadora escribe su trabajo" (Turnitin, 2020). Se adivina un horizonte donde no es descartable que resistirse a los trabajos generados por ChatGPT y similares y detectar cuáles están escritas por máquina sea un ejercicio inútil. Entonces, ¿cómo podemos delimitar el uso de estas nuevas herramientas?

Ya hemos hablado de restringir la asignación en las ponderaciones por puntajes en estos temas. Alternativamente, se podrían establecer tareas escritas reflexivas y contextualizadas que la IA no podría generar. Pero esto requiere mucho tiempo y esfuerzo por los profesores

En todo caso, los profesores podrían y deberían explorar y penetrar con los estudiantes en la ética y los límites de la IA generativa y hacerlo con cuestiones tales como:

- ¿Qué calificación y qué efectos tiene en los conceptos morales y por qué interactuar con un agente experto en palabras que no tiene moral intrínseca ni experiencia del mundo?
- ¿Escribir con ChatGPT equivale a plagio?

Pero más allá de esto hay una cuestión clave. Si restringimos la evaluación a lo que es la lectura y la revisión de proyectos o ensayos en el marco exclusivo de esa lectura, damos la entrada a la impostura porque en ese marco, como dice la revista *Nature* (Else, 2023), es imposible distinguir lo que escribe un *transformer* de lo que escribe un humano. O, en todo caso, la resolución de ese problema nos remite a una futura IA que discerniera que es un autor humano a través del estilo, inflexiones y giros propios y en otros rasgos personales, y en todo caso abundaría en algo que la inteligencia humana le lleva años luz a la IA, como sucede con la psicología cognitiva con respecto al conductismo y al aprendizaje por refuerzo: el análisis de la experiencia, la atribución de sentido través de la expresión escrita y la interacción que, de ella, se pueda producir, como lo son las que se producen en la revisión y supervisión.

Se hace pues hoy más necesaria que nunca una visión de lo que ha de ser una evaluación centrada en el aprendizaje, en los efectos que produce, en cuales son deseables y, sobre todo, en cuáles de sus manifestaciones presentan una demostración de que el fenómeno de aprender (comprender, atribuir sentido, incorporar, ejecutar autónomamente y transferir) se ha producido. Y hacerlo en el transcurso del proceso. En definitiva, hoy es más urgente que nunca contar con una buena evaluación educativa centrada en el aprendizaje y formativa.

## 2. Evaluación formativa. La auténtica evaluación educativa, la de calidad, es inmune al ChatGPT

### La auténtica evaluación educativa, la de calidad, es inmune al ChatGPT

La evaluación es un complejo de procedimientos que, en condiciones normales, se distribuyen a lo largo de todo el proceso educativo a través de distintos recursos y procedimientos y difícilmente, para que sea una evaluación de calidad, se hace en un único momento o en unos pocos momentos a lo largo del curso, o utilizando una única técnica o instrumento. Si se hace así, como se hacía, y aún se hace con demasiada frecuencia con la evaluación, fiándola a un examen o a un proyecto, se corre el riesgo de tomar decisiones sesgadas o arbitrarias. Además de crear en los alumnos una ansiedad que los desmotiva y perturba su aprendizaje.

Actualmente tenemos el conocimiento de cómo se produce el aprendizaje y de cómo se realizan las prácticas pedagógicas en las mejores condiciones. Lo hemos obtenido a través de investigaciones que ha aportado resultados y evidencias. También lo tenemos para organizar una evaluación de manera confiable sin necesidad de recurrir exclusivamente a las consabidas pruebas de conocimiento o ensayos, a final de curso, ciclo, trimestre, cuatrimestre o materia. Todo lo que se ha publicado en la literatura, como consecuencia de esas investigaciones y de esa elaboración científica, nos proporciona esos procedimientos y esos métodos. Es papel de los diseñadores de la instrucción en todos sus niveles el conocerlos y el aplicarlos en TODOS LOS ESCENARIOS. Y su responsabilidad el no hacerlo así.

Como en otros casos y en otros procesos, ahora con ChatGPT y antes con la pandemia, se están poniendo de relieve hábitos, tradiciones y procesos que estaban pasando desapercibidos y que han manifestado por un lado su vulnerabilidad, sus insuficiencias y por otro las alternativas que la teorías pedagógicas nos venían suministrando, siendo muchas veces ignoradas, y por otro el papel que la tecnología y las redes digitales con sus *affordances* nos estaban suministrando para la implementación de esos enfoque teóricos así como la oportunidad de validarlos con la práctica.

Con respecto a la evaluación, la segunda cuestión clave que se ha de tener en cuenta es que desde siempre la evaluación ha sido un proceso que, para ser de calidad, difícilmente se puede desvincular del **quehacer cotidiano del alumno**. Pues bien, si siempre ha sido así, ahora es más vital ese factor: Ahora con la amenaza de ChatGPT y Transformers IA y antes con la pandemia, cuando no hay una interacción directa entre alumnos y profesores, es cuando debemos ser conscientes de su importancia, de que **esa interacción debe ser especialmente detallada, en su práctica y programación, así como frecuente**, de forma casi continua, en la medida en que las condiciones físicas de presencia directa mediante entrevistas y supervisiones o a través de la tecnología, lo permitan. Alguien planteará que eso es imposible o muy difícil. Hoy no es así. Y no por la eficiencia de las herramientas tecnológicas, que también, sino porque hay metodologías docentes que lo permiten. Desde hace ya tiempo. Y entre ellas, como después veremos, la principal es la de *Mastery learning* (Zapata-Ros, 2018a).

Se trata pues de una cuestión de metodología docente, de cómo los profesores hacemos las cosas.

La primera conclusión que he sacado, en toda mi experiencia, es que la evaluación e incluso la acreditación de conocimientos o de habilidades no debe abordarse como algo aislado, condicionado exclusivamente por variables limitadas a la coyuntura o por otros factores recientes y directos. Si lo hacemos así se pueden derivar más males que beneficios, y además los resultados van a ser irreparables.

La evaluación en todos sus aspectos debe de estar, en todas sus facetas y ejecuciones, integrada en un proceso de diseño instruccional amplio y detallado lo más posible.

Pero ¿qué es el diseño instruccional? Un diseño es, en general, un proyecto o un plan que configura algo antes de su realización o de su desarrollo. Pero que también permite saber en qué punto de él estamos en su ejecución, y cómo relacionar lo que estamos haciendo con los resultados últimos de lo que hagamos. Pudiendo retomar y modificar decisiones, en progreso, para cambiar la ejecución y que ésta se ajuste con nuestros objetivos en un *feedback* continuo.

Los procesos y programas formativos, y la educación, también se diseñan.

En general el *diseño instruccional*, como veremos, se define como "un proceso sistemático que se emplea para desarrollar programas de educación y capacitación de manera continua y confiable".

Existe una amplia base de modelos de diseño instruccional y de desarrollos teóricos considerados como marcos para el desarrollo de módulos formativos o clases que aumenten, haciéndolo de acuerdo con sus pautas y preceptos, la posibilidad de aprender, de fomentar la participación de los alumnos para que aprendan más rápido, y obtener de esta forma niveles más profundos de entendimiento.

Los principales creadores e impulsores del diseño instruccional y de sus bases teóricas ha sido David Merrill (2002 y 2012) y Charles M. Reigeluth.

Así la Teoría del Diseño Instruccional clásica de Reigeluth se considera, en general, como un instrumento con un doble fin: facilitar el aprendizaje y el desarrollo humano, o mejor dicho facilitar el desarrollo humano en la medida que se consigue un mejor aprendizaje. Es en esencia una teoría situacional. Sostiene que los métodos y situaciones de aprendizaje son esenciales para que el aprendizaje tenga lugar de forma efectiva. Está constituida por una serie de principios para organizar la enseñanza en un esquema complejo de elementos más pequeños, y por tanto más cerca de la comprensión individual, que posteriormente son insertados en estructuras conceptuales interrelacionadas.

Los métodos del diseño instruccional suponen un ciclo continuo y una evaluación formativa (Zapata-Ros, 2013) que permiten introducir mejoras sobre el proceso en el diseño del programa educativo, sin necesidad de concluir.

El diseño instruccional permite pues partir, en un proceso inverso, desde la consideración de *cuáles son los resultados deseados del proceso de aprendizaje* para organizar las pautas, para distribuir en el tiempo todo lo demás, incluida la secuencia de acciones y de contenidos. Permite distribuir no solo la entrega de materiales, de actividades y de ayuda, sino de establecer cuáles son los logros parciales que, en una cadena, lleven al alumno a conseguir los logros de la unidad didáctica, el módulo o la materia. De forma que su consecución sea gradual, continua, sin saltos.

Hemos dicho que el diseño instruccional consiste también en la aplicación de los principios teóricos que se conocen sobre como los individuos aprenden para organizar la educación. Veamos particularmente cómo funciona esto, la aplicación de los principios teóricos, en un punto concreto, el de la evaluación, que es el tema de este post. Lo hace a través del *principio de demostración*.

## **La evaluación y el Principio de Demostración**

Pero antes expliquemos un poco más qué son los principios del aprendizaje. Como referencia teórica, hemos utilizado, en todo lo que hemos escrito aquí y en material que he ofrecido en estas circunstancias para ayudar al diseño de enseñanza *online* y para la evaluación, el trabajo de David Merrill (2002 y 2012) conocido como los "Primeros [o principales] principios de la instrucción" (*First principles of instruction*, Prentice Hall). Todo ello está desarrollado en el capítulo 2 y se puede ampliar allí con más detalle y en las referencias que contienen enlaces a los documentos originales.

Estos principios son una constante de la educación en todas las investigaciones, son aceptados por todos los teóricos, hay un consenso sobre ellos.

Merrill estableció cinco principios: El aprendizaje ha de ser centrado en el problema, el principio de activación, el de demostración, el de aplicación y el principio de integración o de transferencia.

El principio que fundamenta la evaluación es el principio de demostración. Y es el que debiéramos tener presente en todo momento en que hablásemos de ella. Simplificando consiste en que el alumno ha de aportar evidencias de su aprendizaje, en un sentido pedagógico, es decir ha de demostrar que comprende que atribuye sentido, que aplica las habilidades conseguidas de forma autónoma y las transfiere a contextos distintos.

¿Qué dice?

En primer lugar, y para entender lo que sigue, hemos de tener en cuenta que los principios de la instrucción están enunciados en forma de criterios que ha de cumplir la enseñanza para que el aprendizaje sea eficiente, de calidad. Hemos de recordar que la evaluación forma parte de la **instrucción**.

### ***Principio 3-Demostración.-***

El aprendizaje se promueve cuando la instrucción demuestra lo que se debe aprender en lugar de simplemente decir información sobre lo que se debe aprender.

Así la demostración es un factor de consistencia: El aprendizaje se promueve cuando la demostración es coherente con el objetivo de aprendizaje, con:

- (a) ejemplos y contraejemplos de conceptos,
- (b) demostraciones para procedimientos,
- (c) visualización (evidenciar visualmente) para los procesos, y
- (d) el modelado (presentación de modelos) para adquisición de valores y conductas

Obviamente ninguna de estas exigencias de la demostración se cumple con el ChatGPT ni el propio GPT las demuestra.

Presentar un informe o un proyecto, por sí mismo no demuestra nada.

La repercusión del diseño instruccional en la evaluación tiene otro punto clave: La cuestión inmaterial, pero sobre la que gira la docencia, el entregar ayuda y contenidos cuando el alumno esté receptivo y los perciba como útiles, estriba en distribuir los logros parciales y evaluar su consecución, y únicamente entonces proseguir. En esto juega un papel central el diseño y, sobre todo, si finalmente los hay, el papel de los Teaching Assistant (TA). Más abajo describiremos como ejemplo lo que hicimos en el curso investigativo de Alcalá.

### **La confianza**

Por último, se trata también de una cuestión de confianza. En una educación de estándares, a la que estamos acostumbrados y donde el ChatGPT, los proyectos a los que sustituye y los exámenes, se han establecido medidas iguales para todos, los estándares, que en su consecución acreditan al alumno para superar el curso o la asignatura. Es un sistema basado en estándares iguales para todos, de superación de pruebas y de exámenes iguales para todos.

Esta dinámica permite trasladar la eficiencia al escrutinio de las pruebas y de sus resultados por padres, familias e inspección, donde además con esa supuesta objetividad los docentes se sienten seguros, aunque por contrapartida lo que se hace es sembrar la semilla de la desconfianza, de que no haya equidad en el proceso. Esto es un caldo abonado para la

industria del fraude y la del ChatGPT como impostor del alumno para demostrar capacidades y aprendizajes en general.

Este hecho viene a confirmar algo que planteamos desde hace tiempo, la necesidad de una nueva educación, la basada en logros personales. En ese marco el entorno social del alumno debe depositar la confianza en el modelo, personalizado a través de sistemas que como el de *Mastery learnig* otorge en los maestros la confianza en valorar que se alcanza el dominio en el aprendizaje. De confiar en el maestro y en su *expertise* docente, de igual forma como la deposita en el médico, en su capacidad de diagnóstico y de decidir un tratamiento adecuado.

### **La evaluación formativa**

En lo que sigue haremos un repaso de lo que se entiende por evaluación formativa, de cómo la podemos relacionar con lo que está sucediendo con ChatGPT y lo que hemos dicho sobre él. Es conveniente que todo ello lo tengamos en mente cuando veamos todo esto.

También veremos otros conceptos y otras formas de evaluación. Por último, haremos un sucinto repaso de actividades que tienen un fuerte carácter de evaluación formativa y que se sugieren como algo que proporciona información sobre el estudiante y sobre su progreso en el camino hacia aprender y obtener el dominio en las destrezas que se pretende que consiga. Pero no nos quedaremos en el mero enunciado o la sugerencia de actividades, veremos qué procesos y que logros son los más importantes en ese camino.

Leer sólo lo que proporciona un informe o un ensayo, prescindiendo de todo lo que estamos viendo y de lo que en general se sabe y se ha investigado, para evaluar a un estudiante, aunque sepamos mucho sobre, él puede ser injusto por insuficiente. En definitiva, evaluar teniendo en cuenta únicamente el texto, en el que como hemos visto, ahora con Chat GPT, es imposible discernir entre un humano y un programa *transformer*, nos lleva al fracaso y al malestar por vernos, como profesor4es, indefensos y desamparados en ese camino. Y lo que, sin ser tan grave es muy negativo, a perder la confianza en la tecnología y en una visión positiva de ella como apoyo nuevo e imprescindible para muchos aprendizajes.

### **Definiciones y validez de la evaluación formativa**

Un trabajo básico pero muy completo es el que Terry Crooks (2001) hizo para la Unidad de Investigación en Evaluación Educativa en la Universidad de Otago (Dunedin, Nueva Zelanda), presentado en la Conferencia Anual de la Asociación Británica de Investigación Educativa, Universidad de Leeds, 13-15 de septiembre de 2001, en la sección de Educación online, titulado *The Validity of Formative Assessments*. Así Crooks (2001) define evaluación como “Cualquier proceso que proporcione información sobre el pensamiento, los logros o el progreso de los estudiantes”.

Para ello se valió de un importante concepto previo, el de validez Utilizó el trabajo y definición de Messick (1989, p.13 Vigency) “La validez es un juicio evaluativo integrado del grado en que la evidencia empírica y los fundamentos teóricos apoyan la idoneidad y pertinencia de las inferencias y acciones basadas en los puntajes de las pruebas u otros modos de evaluación”.

Frecuentemente se utiliza el concepto de evaluación sumativa, más acorde con lo que habitualmente se hace y cómo se entiende la evaluación

De esta forma, la evaluación sumativa tiene la intención de resumir el logro del estudiante en un momento particular, mientras que la evaluación formativa tiene como objetivo promover una mayor mejora del logro del estudiante, además de asegurar, obviamente, lo anterior.

Es decir, se trata de contraponer *evaluación de lo aprendido* y *evaluación para aprender* (*Assessment OF learning versus assessment FOR learning*) en el sentido más simple de las dos expresiones:



Veamos ahora en este contexto algunas de las definiciones que hemos encontrado más precisas de evaluación formativa

Para el Ministerio de Educación, Nueva Zelanda, 1994 (*Assessment: policy to practice*) se trata de

Una gama de procedimientos de evaluación formales e informales (por ejemplo, el seguimiento del desarrollo de la escritura de los niños, registros de hechos concretos y observaciones [igual puede generalizarse a procedimientos seguidos en otros niveles y etapas]<sup>4</sup>) realizados por los profesores en el aula [o en la actividad en línea] como parte integral del proceso normal de enseñanza y aprendizaje para modificar y mejorar el aprendizaje comprensión.

Para Bronwen Cowie & Beverley Bell (1999) (*A model of formative assessment in science education*) es “el proceso utilizado por profesores y estudiantes para reconocer y dar respuesta al proceso de aprendizaje de los estudiantes con el fin de mejorar ese aprendizaje durante el aprendizaje”.

Para Black & Wiliam (1998a, p. 2) (*Inside the black box*) la

Evaluación se refiere a todas aquellas actividades realizadas por los profesores, y por los estudiantes en la autoevaluación, que proporcionan información que se utilizará como retroalimentación para modificar las actividades de enseñanza y aprendizaje en las que participan. Dicha evaluación se convierte en "evaluación formativa" cuando la evidencia obtenida se utiliza en la práctica docente para adaptar la enseñanza a las necesidades detectadas.

Pero quizá la más completa es la definición que hacen Black & Wiliam (2009), de evaluación formativa, acompañada de otra sobre lo que es instrucción.

Para lo primero se basa en sus definiciones (Black y Wiliam, 1998b) y la definición del Grupo de Reforma de la Evaluación (ARG, 2002)

La práctica en el aula es formativa en la medida en que la evidencia sobre el rendimiento del estudiante sea suscitada, interpretada y utilizada por los profesores, los alumnos o sus compañeros, para tomar decisiones sobre los próximos pasos en la instrucción. Próximos pasos que probablemente sean mejores, o mejor fundamentados, que las decisiones que se tomarían se han tomado en ausencia de la evidencia que de esa forma se ha obtenido.

Pero ¿qué es instrucción e instruccional? Aquí y en varios lugares más utilizamos este término. Procede pues definirlo:

---

4 El añadido entre corchetes es mío.

En gran parte del mundo académico, este término tiene una connotación de formación, o de enfoques didácticos de la enseñanza. Sin embargo, en otros y de forma mayoritariamente aceptado, el término "instrucción" significa la combinación de enseñanza y aprendizaje, y es este último sentido el que utilizaremos aquí. En esta definición, "instrucción" se refiere a cualquier actividad que vaya destinada a crear aprendizaje. En España el término equivalente sería *pedagogía*. Pero actualmente también, además de que el enfoque tradicional tiene inconvenientes al sugerir una visión más limitada de la práctica de la enseñanza, no se ajusta demasiado bien a términos y conceptos de la sociedad digital y del conocimiento. Por ejemplo, parecería extraño referirse a los estudiantes como 'recursos pedagógicos unos para otros', parece que hablamos de textos, pizarras, retroproyector, etc. No así si decimos "alumnos como recursos instruccionales unos para otros".

A esto podemos añadir una particular aportación de quien suscribe (Zapata-Ros, 2012)

La evaluación formativa o de proceso se exige para entender no sólo si un alumno ha conseguido aprendizaje o no, sino por qué. Comprender las razones para el éxito o para el fracaso depende de la profundidad del conocimiento que existe de lo que se ha hecho o se está haciendo, y también de la relación apropiada entre el uso de la tecnología, las actividades que se han hecho con ella y el área de conocimiento. Entendiendo como tal la metodología docente para el área (las ideas pedagógicas) y los conocimientos propios de la materia.

Por lo tanto, podemos concluir que, aunque el desarrollo de la evaluación formativa no es la única manera, o incluso la mejor manera, para abrir una gama más amplia de cambios deseables en el aprendizaje en el aula, puede ser particularmente eficaz, en parte porque la calidad de la retroalimentación interactiva es un factor crítico característica para determinar la calidad de la actividad de aprendizaje y, por lo tanto, es una característica central de pedagogía.

Después de esto creemos que el lector se habrá formado una idea cabal y actualizada de lo que es la evaluación formativa. Pero no es sólo importante saber qué es sino cómo se practica: las actividades.

En base a la literatura, las prácticas, en condiciones normales de trabajo en aula y online, comunicadas por docentes y en base a la propia experiencia en la docencia de secundaria, grado y postgrado y formación del profesorado en activo, también en condiciones normales online, podemos establecer una clasificación en los principales tipos de actividad que señalamos.

Resumiendo lo consultado sobre evaluación formativa podemos ver que hay ocho tipos principales de actividades o de prácticas docentes, sugeridos por evidencias (limitadas a la propia experiencia) de su eficacia potencial, y desarrollado con y por profesores en condiciones normales de trabajo en el aula, también sugeridos por la literatura sobre el tema (Wiliam, 2000, Black et al., 2003; Wiliam, 2007):

- Compartir criterios de éxito con los alumnos
- Preguntar en el aula, o en el chat, o verbalmente en Zoom o similar, con anotaciones.
- Marcado de sólo comentarios
- Autoevaluación, con criterios explícitos
- Uso formativo de pruebas sumativas (muy importante).
- Revisiones, con observaciones y marcado de comentarios
- Supervisiones de los alumnos, con reuniones presenciales, videorreuniones o sesiones de videollamada.

- Intervenciones en los foros con comentarios personales, otorgamiento de relevancia en relación con la formatividad y con relación a los objetivos de aprendizaje, a los logros esperados y a las habilidades desarrolladas, sobre todo conceptuales, en este caso.

De estos ocho, cinco se proponen y están estudiadas en *Developing the theory of formative assessment* por Paul Black, King's College London, and Dylan Wiliam, Institute of Education, University of London.

Objeto aparte de estudio y de desarrollo puede ser el tratamiento de cada uno de estos tipos de actividades.

Particular importancia tiene la evaluación de marcado sólo con comentarios (pruebas comentadas) y la evaluación formativa a partir de la evaluación sumativa (ver *Using comment only marking in a mathematics classroom in the Republic of Ireland: experience and learning of a student teacher*, por O' Mahoney and Heinz (2016).

La *calificación de sólo comentario* es un método de evaluación formativa. Implica que el profesor brinde retroalimentación constructiva y metas parciales vinculadas a los objetivos para alentar y apoyar a los estudiantes a desarrollar los objetivos de aprendizaje establecidos independientemente y sus habilidades de autoevaluación.

Este tipo de evaluación se considera una estrategia clave de aprendizaje, también, y sobre todo, en el aula.

Como resultado, la información proporcionada por esta evaluación puede utilizarse como retroalimentación para influir y modificar las actividades de enseñanza y aprendizaje. Según Petty (2009), se puede utilizar esta forma de evaluación para motivar y animar al alumno al mismo tiempo que mejora él y mejora su trabajo. Hay dos formas principales de evaluación, evaluación sumativa y evaluación formativa.

Sobre lo segundo, las evaluaciones acumulativas se dan periódicamente para determinar la comprensión de un estudiante del tema en un momento particular en relación con los estándares de contenido. La frecuencia aumenta proporcionalmente la comprensión y la adquisición, debido al dominio en aspectos parciales que aumenta la capacidad para aprender lo siguiente.

El marcado de sólo comentarios y la relevancia suficiente de intervenciones en foros y chats tienen, aparte de lo visto, una importancia determinante y clave en la metodología docente y de evaluación conocida como *mastery learning*, como podemos ver en otros documentos.

Cuestiones muy importantes que nos planteamos a propósito de esta modalidad de evaluación, sobre todo en su práctica:

*¿Hasta cuándo haremos marcado de comentarios, durante qué tiempo o cuántas rondas haremos? ¿en cuántas revisiones? ¿cómo sustituiremos la nota, la puntuación numérica clásica, por la calificación, obviamente textual, para señalar la superación del módulo o de la unidad de contenidos o la consecución [dominio] de la habilidad que tratemos?*

A todas estas preguntas se da respuesta en la modalidad de "dominio del aprendizaje" (*Mastery learning*). Modalidad en la que sucintamente se dice que sólo se pasara a la parte siguiente cuando se consiga un dominio de la actual. Obviamente, en este contexto tienen un papel esencial los *Teaching Assistant*.

Hay otra cuestión que también es clave: Este método, esta modalidad de evaluación, supone la confianza y la conformidad en la capacidad del docente, su *auctoritas*, para determinar cuándo se ha alcanzado el dominio. No en balde se llama maestría, evocando la relación entre maestro y aprendiz que se establecía en otras etapas de la historia de la educación y de la instrucción.

La evaluación sumativa es la que sólo proporciona los criterios de promoción del alumno al final del módulo, la asignatura o el curso.

Por tanto, hay básicamente dos formas de evaluación, evaluación sumativa y evaluación formativa.

Las evaluaciones acumulativas se ofrecen periódicamente para determinar el grado la comprensión del tema de un estudiante en un momento particular en relación con los estándares de contenido. Otra cuestión es si coincide con la evaluación del aprendizaje. Aprendizaje no siempre supone sólo comprensión, para muchos de nosotros supone dominio.

De acuerdo con Garrison y Ehringaus (2007), la evaluación sumativa ocurre demasiado poco y de forma distanciada en el camino de aprendizaje del estudiante para que pueda producir ajustes instructivos durante el proceso de aprendizaje.

La evaluación formativa ocurre cuando los resultados de la evaluación se consideran y se utilizan para modificar la enseñanza para satisfacer las necesidades de los estudiantes casi instantáneamente. También se conoce como evaluación para el aprendizaje (Black y William, 1998). La evaluación formativa es útil tanto para estudiantes como para profesores. Proporciona retroalimentación a los docentes sobre la efectividad de sus estrategias y enfoques de enseñanza, que luego pueden ajustar para mejorar el aprendizaje de los estudiantes. En segundo lugar, la evaluación formativa puede ser útil para los estudiantes ya que la retroalimentación es enfocada y precisa para que tengan la capacidad de identificar áreas específicas que necesitan trabajar y si necesitan ajustar su forma de pensar (Phelan et al., 2011).

Por último, más allá de definiciones, conceptos y actividades, conviene señalar el sentido de los procesos y de otros aspectos de la evaluación formativa en un contexto más amplio.

Para proporcionar una base amplia de la evaluación formativa, tres procesos son centrales: establecer dónde se encuentran los estudiantes en su aprendizaje, establecer hacia dónde se dirigen y establecer lo que se debe hacer para lograrlo y cómo ayudarlos a llegar allí (Thompson y William, 2007 a partir de Ramaprasad (1983)). Con referencia a estos tres procesos estableceremos los principales aspectos a considerar en la evaluación formativa en todos los casos:

	A dónde va el alumno	Dónde está el alumno ahora mismo	Cómo llegar allí
El profesor	<b>1</b> Aclarar el aprendizaje: las intenciones y los criterios para el éxito.	<b>2</b> Construir una arquitectura efectiva de la clase: Fomentar discusiones en el aula y proponer tareas de aprendizaje que provoquen la evidencia del estudiante propia de la comprensión.	<b>3</b> Proporcionar comentarios y estímulos que mueva a los alumnos hacia adelante, que hagan que progresen.
El igual	Comprender y compartir intenciones de aprendizaje y criterios para el éxito	<b>4</b> Activar a los estudiantes como recursos educativos entre sí.	
El alumno	Entender el aprendizaje, las intenciones y los criterios para el éxito	<b>5</b> Activar a los estudiantes como dueños de su propio aprendizaje	

En conclusión, un sistema educativo, que implemente estas ideas, es absolutamente inmune al uso de ChatGPT como sustituto total o parcialmente en la demostración del aprendizaje. Otra cuestión es su uso de forma crítica como recurso educativo, con un carácter similar al que tiene actualmente cualquier buscador o enciclopedia digital.

### 3. Una pedagogía para la IA generativa

En esta parte recordaremos que el dominio y la práctica de IA es mucho más amplio que la IA generativa y que, en la educación apoyada por la IA o que la integra. También que hay visiones y orientaciones de la IA educativa distintas, como lo es un nuevo aprendizaje adaptativo o que lo integra. Es el aprendizaje inteligente (Zapata-Ros, 2018b).

No obstante, no se podría ignorar el impacto que actualmente tiene el mundo de los LLM y del ChatGPT. Tampoco los planteamientos que se hacen y a los que, por su alcance, precisan una respuesta. Estos planteamientos son que la nueva IA tenga el potencial de revolucionar los métodos de enseñanza, la evaluación y la ayuda al alumno, existentes. O que la IA tenga la capacidad de crear un sistema educativo completamente nuevo

La modalidad de inteligencia artificial generativa se ha impuesto y avanza rápidamente. Pero hay otra modalidad, como es la basada en la programación declarativa y en la programación lógica, que cayó hace mucho en desuso más próxima a como el pensamiento humano procesa la información.

No sabemos por qué, el devenir económico, cultural y humano de la sociedad actual, entre las dos formas posibles de simular un cerebro humano, como las que expresa Noam Chomsky (2023) en el párrafo siguiente, se ha elegido la de la IA generativa:

*La mente humana no es, como ChatGPT y sus similares, un pesado motor estadístico para la coincidencia de patrones, que se atiborra de cientos de terabytes de datos y extrapola la respuesta conversacional más probable o la respuesta más probable a una pregunta científica. Por el contrario, la mente humana es un sistema sorprendentemente eficiente y hasta elegante que opera con pequeñas cantidades de información; no busca inferir correlaciones brutas entre puntos de datos, sino crear explicaciones.*

*Por ejemplo, un niño pequeño que adquiere un idioma está desarrollando — inconsciente, automática y rápidamente a partir de datos minúsculos— una gramática, un sistema extraordinariamente sofisticado de principios y parámetros lógico.*

Pero ésa ha sido la elección. Y tras décadas de investigación y desarrollos, ahora se ha decidido por parte de Open AI, y después por otras empresas, que este tipo de IA, a la que en lo sucesivo llamaremos para simplificar IA en vez de IA generativa, pase de los laboratorios a la aplicación práctica en la sociedad. No sabemos si ha sido sólo efecto de una potente campaña o porque la sociedad lo ha aceptado así, pero el hecho es que se ha producido un colosal despliegue de esta IA.

La educación no ha sido indiferente a esa aceptación ni a ese despliegue.

En todo caso este hecho inevitable y el enunciado previo nos marcan las coordenadas que nos plantean la necesidad de una nueva Pedagogía: La necesidad y la fluctuación entre los dos polos que hay entre la elección que la sociedad realiza y el presupuesto de Chomsky.

Aun así, si optamos por tomar como referencia esta situación, la educación, hasta donde se ha dicho (Dawson et al., 2023), se enfrenta a dos posibles resultados.

En primer lugar, los que se derivan de que la IA tenga el potencial de revolucionar los métodos de enseñanza, la evaluación y la ayuda al alumno, existentes. De lo primero y lo segundo hablaremos después.

En segundo lugar, según estos mismos autores y lo que parece ser un sentir común, la IA también tiene la capacidad de crear un sistema educativo completamente nuevo. Y se cita como justificación que el pensamiento creativo y la resolución de problemas son fundamentales

en entornos modernos y muy complejos. Y esta IA podría ayudar a los alumnos a enfrentarse a esos problemas.

Estas tres cuestiones ya por sí solas plantean tres desafíos y otros tantos ejes de posibles investigaciones y para su difusión en interesantes contribuciones.

Esto es en sentido positivo, pero hay otras cuestiones que pueden ser planteadas en sentido negativo o como interrogantes:

Así nos planteamos cuáles son los impactos que se producirán cuando la IA comience a servir como un socio activo<sup>5</sup> en acciones sociales, creativas e intelectuales sostenidas a lo largo del tiempo, no sólo puntuales o como repuesta a preguntas aisladas, que es lo que ahora más frecuentemente se hace. En este momento esos impactos en las prácticas, que puedan existir, son desconocidos. Ya hemos dicho que indagar esas prácticas también puede ser objeto de investigación y de eventuales contribuciones.

Otras son preguntas que se abren sobre cómo del impacto de una dependencia excesiva, en los sistemas de IA en la educación, se podría derivar una disminución de muchos de los rasgos que nos hacen humanos. En este caso la pregunta es si ese impacto contribuiría a una disminución de la formación “humana” de los alumnos.

Entre esos rasgos se incluyen, pensando a bote pronto, la autorregulación, la metacognición, la orientación a objetivos, la planificación, las tormentas de ideas creativas y una variedad de habilidades que podrían verse afectadas negativamente por la automatización o la toma de control de las máquinas.

Paradójicamente estos rasgos, a más de ser humanos, son constitutivos de lo que se ha llamado pensamiento computacional (Zapata-Ros, 2015 y 2022; Grover, 2018; Wing, 2006 y 2008), de manera que un uso no controlado podría derivar en una nula o una mala adquisición de ese tipo de pensamiento que hoy se considera constitutivo de una nueva alfabetización en la sociedad digital (Zapata-Ros, 2015 y 2022).

En particular para justificar el papel que tradicionalmente se asigna a la metacognición, la autorregulación y la retroalimentación podemos citar las siguientes dos aportaciones como más significativas:

Hilpert et al. (2023) abordan la necesidad de participación en el aprendizaje de STEM, con una mayor necesidad de autorregulación de los estudiantes. Se basan en una extensa historia de la ciencia de la investigación del aprendizaje que utiliza datos de seguimiento digital para crear construcciones cognitivas que brindan información sobre el compromiso, las redes sociales, la comunidad y la metacognición. En este documento, detallan la importancia de la regularidad de la participación como un fuerte predictor del resultado del curso y los efectos del uso de una intervención de la ciencia de aprender a aprender para fomentar el SRL de los estudiantes y la participación continua. Sus resultados sugieren efectos prometedores y sostenidos de este entrenamiento, lo que plantea la necesidad de considerar enfoques teóricos que integren las observaciones del comportamiento con las construcciones cognitivas en la educación digital.

Bauer et al. (2023) abordan un área crítica de aprendizaje relacionada con la retroalimentación. La retroalimentación es fundamental para guiar el progreso de los estudiantes y los enfoques tradicionales se basan principalmente en la observación humana.

Sin embargo, el mismo autor, señala un posible progreso del que después hablaremos, cuando lo hagamos de aprendizaje inteligente como aprendizaje inclusivo:

*Con el desarrollo de grandes modelos de lenguaje y el procesamiento del lenguaje natural en general, existen nuevas oportunidades para ofrecer retroalimentación a los estudiantes. Detallan cómo los artefactos textuales pueden mejorarse con la retroalimentación*

---

<sup>5</sup> En este punto nos planteamos si cabría decir “servir como socio”, o más bien “ser socio” o, por el contrario, “servir como herramienta”.

*de la IA. Ofrecen un marco para conectar los procesos de retroalimentación con el apoyo estudiantil adaptativo. A medida que el aprendizaje digital crece en importancia en los entornos educativos, la inclusión de artefactos más diversos y multimodales requerirá una actualización similar de la teoría y las construcciones para garantizar la retroalimentación como motor del éxito general de los estudiantes.*

### **Una pedagogía y una teoría del aprendizaje de la IA**

Para abordar estas preguntas que nos hacemos, y en general ante la realidad de un despliegue efectivo de los sistemas de IA generativa en la educación, nos planteamos que es preciso hacerlo desde un punto de vista teórico, más allá de los resultados que sobre las interrogantes señaladas nos proporcione la investigación empírica. Un punto de vista teórico que no guíe y dirija en nuevas rondas, no sólo en esta primera, tanto en la investigación como la práctica.

Ese marco nos proporcionará los asideros y las andaderas para garantizar que los principios, los valores y las construcciones confiables configuren el uso de la IA en la educación. De esta forma nos aseguramos de que los valores, la investigación realizada hasta ese momento, los intereses y preocupaciones de las partes interesadas y las contribuciones que se vayan haciendo sigan integrándose de forma cohesionada en el escenario de la IA en la educación. Al menos como deseo y horizonte.

Como consecuencia, en futuros desarrollos defenderemos la importancia crítica de la teoría en el diseño, desarrollo y despliegue de la IA en la educación. Al hacerlo, nos planteamos de forma igualmente crítica la relevancia y la continuidad de las teorías de aprendizaje ya existentes, cuando la IA se constituya como una realidad en las aulas.

Así, hacemos una llamada a considerar nuevos marcos, modelos y formas de pensar. A aquellos que incluyen la presencia de agentes no humanos. Es decir, a los que incluyan una nueva tecnología, que se parece más a un socio activo que a una tecnología simple, como sucedía hasta ahora, incluso con la tecnología digital, que respondía más a interrogantes lógicas simples que a un agente con capacidad para tomar decisiones intermedias sin intervención humana, del alumno o del profesor. En ese sentido cabe ahora hablar de un trinomio, más que de un binomio, el formado por el profesor, el alumno y la IA generativa.

Es precisamente eso lo que genera también una serie de preguntas importantes para considerar en elaboraciones teóricas futuras, precisamente sobre la revisión de las teorías de aprendizaje basadas en las configuraciones existentes. Y cuáles serían en este caso sus alternativas.

Ya hemos señalado que el uso de la teoría en la investigación educativa constituye la lente, el filtro, a través del que vemos y explicamos lo que observamos (Biesta et al., 2011). ¿Pero es útil la lente que tenemos ahora?

Las teorías existentes se establecieron y se derivaron de un contexto educativo distinto al de los enfoques de enseñanza emergentes y al de los sistemas técnicamente establecidos. Es de prever más aún que, a medida que la adopción de la IA impacte en la enseñanza y el aprendizaje, se produzcan dudas acerca de la idoneidad en continuar con las teorías de aprendizaje existentes.

Sin embargo, como hemos dicho, la integración del desarrollo e implementación de sistemas educativos basados en IA en marcos teóricos adecuados es esencial para avanzar y lograr resultados de aprendizaje óptimos. En base a estas ideas, planteamos las siguientes interrogantes:

- ¿La integración de la IA en la educación requiere modificaciones o revisiones en la forma en que aprendemos?

- ¿O, por el contrario, se requiere una reestructuración completa, lo que conlleva la necesidad de nuevas teorías?
- ¿Qué debería ofrecer la teoría a los educadores cuando se incluya la IA?

Con relación al modo en que se persiguen y abordan estas cuestiones, nos planteamos: En los programas LLM, la facilidad de acceso a los contenidos y la capacidad de producirlos, ¿disminuyen la importancia del conocimiento declarativo?

Se ha hablado mucho hasta ahora del constructivismo social vinculado al aprendizaje apoyado por las tecnologías digitales y sociales. De hecho, es el paradigma imperante durante mucho tiempo, con el que todos se sentían satisfechos, incluso impulsaba reformas de gran calado como fue la reforma de Bolonia en la Educación Superior europea. Pues bien, del mismo modo nos podríamos plantear si los puntos de vista existentes sobre el constructivismo social ¿avalan la IA como un agente social en el aprendizaje?

También podríamos hacer incursiones microteóricas. Como, por ejemplo, plantearnos si los corpus teórico-prácticos existentes sirven suficientemente, o son los adecuados, para la inclusión de la IA generativa en las aulas o es necesario actualizarlos o incluso repensarlos por completo.

### **El aprendizaje inteligente**

Todo esto podría servir para orientar el debate y la investigación en torno a la nueva visión que se abre con la IA generativa. Pero hay unos planteamientos anteriores que ya hemos citado y que no conviene olvidar: Es la consideración del llamado aprendizaje inteligente (Zapata-Ros, 2018b; Hwang et al, 2008) como una fase superior del aprendizaje adaptativo.

El constructo de "entornos de aprendizaje inteligentes" puede considerarse, a partir de la definición de Hwang et al (2008), como el de "entornos de aprendizaje apoyados por la tecnología que realizan adaptaciones y brindan el soporte adecuado con la tecnología de detección y de recomendación" (por ejemplo, orientación, retroalimentación, consejos o herramientas a utilizar) en los lugares adecuados y en el momento oportuno, según las necesidades individuales de cada alumno, determinados mediante el análisis de sus comportamientos de aprendizaje, su trayectoria de rendimiento y los contextos en línea y del mundo real en los que se encuentran, tanto grupales y sociales como compuesto por los recursos de que se provee o a los que accede, descarga, lee o elabora.

Y aprendizaje inteligente podría considerarse consecuentemente (Zapata-Ros, 2018) como aquel aprendizaje que pueden desencadenar y producir los sistemas de aprendizaje inteligente, y que se puede entender de forma diferenciada por las funciones y efectos que no se pueden realizar sólo en ambientes de aprendizaje adaptativo y de aprendizaje contextualizado, pero sin excluirlas. El aprendizaje social y el aprendizaje colaborativo de esta forma se puede entender como una subcategoría del aprendizaje contextualizado.

Resumiendo, se puede decir pues que el aprendizaje inteligente es el aprendizaje adaptativo y contextualizado con mecanismos de detección, de respuesta y de recomendación.

Entonces Zapata-Ros (2018) se planteaba

La presencia y el avance de las tecnologías de detección y de recomendación están ofreciendo grandes y determinantes oportunidades para desarrollar nuevos entornos de aprendizaje. Definitivamente el aprendizaje inteligente es un concepto extremadamente innovador, como en otros ámbitos, para desarrollar visiones y planteamientos distintos. En este caso para crear entornos de aprendizaje más potentes y útiles que los enfoques de aprendizaje ayudados por la tecnología ya existentes, mediante la incorporación de nuevas herramientas y nuevos criterios para el aprendizaje.

La cuestión ahora sería: Cuál es el estado ahora de aquello que se abrió como un desarrollo del aprendizaje adaptativo y cuyas primeras y muy contadas experiencias, en relación con la Educación Superior, se vieron en el trabajo citado (Zapata-Ros, 2018). Qué experiencias y qué conclusiones se han obtenido de las investigaciones en esta línea, no sólo en la Educación Superior, si es que las ha habido. Y finalmente cuál es su estado y cuáles son sus perspectivas en relación con la IA de nuevo cuño, la de los sistemas generativos.

Con todo lo dicho se abre un espacio, solo iniciado de forma tímida e incipiente en relación con la relevancia que se le atribuye a este fenómeno, a posibles trabajos que esencialmente incluyan investigaciones empíricas y/o el desarrollo y prueba de tecnologías en relación con los LLM generativos, estudios de casos o desarrollos teóricos sobre aprendizaje con IA, ChatGPT y sistemas transformadores LLM y Educación. Y también a desarrollos y experiencias sobre aprendizaje inteligente y evaluación formativa en contextos de IA.

Tras ellos la creación de una teoría y de una pedagogía de la IA educativa que vaya más allá de las exégesis y de las interpretaciones.

## REFERENCES

- ARG. (2002). *Assessment for Learning: 10 principles*. Available on the Assessment Reform Group website : [www.assessment-reform-group.org.uk](http://www.assessment-reform-group.org.uk).
- Barrow, B. (2019). *Bayesian Analysis in Natural Language Processing: Cohen*. <https://linguistlist.org/issues/30/30-1843/>
- Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., Fischer, M. R., Fischer, F. (2023). Uso del procesamiento del lenguaje natural para respaldar la retroalimentación entre pares en la era de la inteligencia artificial: un marco interdisciplinario y una agenda de investigación. *Revista británica de tecnología educativa*. <https://doi.org/10.1111/bjet.13336>
- Biesta, G., Allan, J., & Edwards, R. (2011). The theory question in research capacity building in education: Towards an agenda for research and practice. *British Journal of Educational Studies*, 59(3), 225-239.
- Black, E. (1995). Behaviorism as a learning theory. Retrieved from <http://129.7.160.115/inst5931/Behaviorism.html>
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-149.
- Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British educational research journal*, 29(5), 623-637.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31.
- Blanco-González, A., et al. (2022). Preimpresión en arXiv <https://doi.org/10.48550/arXiv.2212.08104>.
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, 13(6), 4-16. <https://www.jstor.org/stable/pdf/1175554.pdf>
- Bloom, B. (1968). Learning for Mastery. Instruction and Curriculum. <https://files.eric.ed.gov/fulltext/ED053419.pdf>
- Bloom, B. S. (1976). *Human characteristics and school learning*. McGraw-Hill.

- Bloom, B. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as effective as One-to-One Tutoring, *Educational Researcher*, 13:6 (4-16).
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P. ..., & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901
- Clark, D. (2023) Twitter. <https://twitter.com/DonaldClark/status/1618422817590243333>
- Bundy, A. (2007). Computational thinking is pervasive. *J. Scient. Pract. Comput.*, 1, 67–69.
- CIT (Center for Instructional Technology). (2013). Building a Coursera Course Version 2.0 [https://docs.google.com/document/d/1ST44i6fjoaRHvs5IWYXqJbiI31muJii\\_iqueJ\\_y1pxG0/edit?pli=1](https://docs.google.com/document/d/1ST44i6fjoaRHvs5IWYXqJbiI31muJii_iqueJ_y1pxG0/edit?pli=1)
- Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam Chomsky: The False Promise of ChatGPT. *The New York Times*, 8.
- Clark, D. (2023). *OpenAI releases massive wave of innovation.* <https://donaldclarkplanb.blogspot.com/2023/02/>
- Clark, D. (2023b). <https://twitter.com/DonaldClark/status/1618422817590243333>
- Cohen, S. (2019). Bayesian analysis in natural language processing. *Synthesis Lectures on Human Language Technologies*, 12(1), 1-343.
- Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles, Policy & Practice*, 6(1), 101-116.
- Crooks, T. (2001). The validity of formative assessments. In *British Educational Research Association Annual Conference, University of Leeds* (pp. 13-15). <http://www.leeds.ac.uk/educol/documents/00001862.htm> (en documentos [http://www.leeds.ac.uk/bei/Education-line/browse/all\\_items/119155.html](http://www.leeds.ac.uk/bei/Education-line/browse/all_items/119155.html))
- Crosslin, M. (2014). Designing a Dual Layer cMOOC/xMOOC. <http://www.edugeekjournal.com/2014/05/04/designing-a-dual-layer-cmoocxmoooc>
- Dawson, S., Joksimovic, S., Mills, C., Gašević, D., & Siemens, G. (2023). Advancing theory in the age of artificial intelligence. *British Journal of Educational Technology*.
- Driscoll, M. (1994). Gagne's theory of instruction. *Psychology of Learning for Instruction. Boston, MA, Allyn and Bacon*, 329-358. <https://psycnet.apa.org/record/1993-99148-000> [http://ocw.metu.edu.tr/pluginfile.php/9013/mod\\_resource/content/1/driscoll-ch10%20\(1\).pdf](http://ocw.metu.edu.tr/pluginfile.php/9013/mod_resource/content/1/driscoll-ch10%20(1).pdf)
- Driscoll, M. (2000). *Psychology of learning for instruction*. Needham Heights, MA: Allyn & Bacon.
- Drury, B. (2019). Bayesian Analysis in Natural Language Processing, in *Review: Computational Linguistics; Text/Corpus Linguistics: Cohen (2019)*. <https://linguistlist.org/issues/30/30-4380/>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature*, 613(7944), 423-423.
- Fink, L. D. (2003). *A Self-Directed Guide to Designing Courses for Significant Learning*. <http://www.deefinkandassociates.com/GuidetoCourseDesignAug05.pdf>
- Gagné, R. M. (1965). *The conditions of learning and theory of instruction* (1st ed.). New York, NY: Holt, Rinehart & Winston.
- Gagné, R. M., & Briggs, L. J. (1974). *The principles of instructional design* (1st ed.). New York, NY: Holt.
- Gagné, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York, NY: Holt, Rinehart & Winston.
- Gagne, R. M., & Medsker, K. L. (1996). *The conditions of learning: Training applications*. Fort Worth, TX: Harcourt Brace College Publishers.
- Garrison, C., & Ehringhaus, M. (2007). Formative and summative assessments in the classroom.

- Gao, C. A. *et al.* (2022). Preimpresión en bioRxiv <https://doi.org/10.1101/2022.12.23.521610>
- Garcez, A. D. A., & Lamb, L. C. (2020). IA neurosimbólica: la tercera ola. *preimpresión de arXiv arXiv:2012.05876*
- Graves, A. (2012). Sequence transduction with recurrent neural networks. *arXiv preprint arXiv:1211.3711*. <https://arxiv.org/abs/1211.3711> Graves, 2012
- Graves, A. (2012). Sequence transduction with recurrent neural networks. *arXiv preprint arXiv:1211.3711*. <https://arxiv.org/abs/1211.3711>
- Grove, J. (2023). The Chat GPT revolution of academic research has begun. *Times Higher Education*.
- Grover, S. (2018, March 13). *The 5th 'C' of 21st century skills? Try computational thinking (not coding)*. Retrieved from EdSurge News: <https://edtechbooks.org/-Pz>
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, GZ (2019). XAI—Inteligencia artificial explicable. *Ciencia Robótica*, 4 (37), eay7120.
- Hilpert, J. C., Greene, J. A., & Bernacki, M. (2023). Leveraging complexity frameworks to refine theories of engagement: Advancing self-regulated learning in the age of artificial intelligence. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13340>
- Kruse, O. (2006). The origins of writing in the disciplines: Traditions of seminar writing and the Humboldtian ideal of the research university. *Written Communication*, 23(3), 331–352.
- Langlais, P.C. (2023). *ChatGPT: comment ça marche?* *Sciences Communes*, Blog de Hypotheses, <https://scoms.hypotheses.org/1059>
- Merrill, M. D. (1991). Constructivism and instructional design. *Educational technology*, 31(5), 45-53.
- Merrill, M. D. (2002a). First principles of instruction. *Educational technology research and development*, 50(3), 43-59.
- Merrill, M. D. (2002b). First principles of instruction. *Educational Technology Research and Development*, 50(3), 43-59. *Instructional-Design Theories and Models, Volume III: Building a Common Knowledge Base*.
- Merrill, M. D. (2009). *First Principles of Instruction*. In C. M. Reigeluth & A. Carr (Eds.), *Instructional Design Theories and Models: Building a Common Knowledge Base (Vol. III)*. New York: Routledge Publishers.
- Merrill, M. D. (2012). *First principles of instruction*. John Wiley & Sons.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: Macmillan.
- Ministry of Education, New Zealand. (1994). *Assessment: policy to practice*. Wellington, New Zealand: Learning Media.
- Newton, P. M. (2018). How common is commercial contract cheating in higher education and is it increasing? A systematic review. *Frontiers in Education*, 3(67), <https://doi.org/10.3389/educ.2018.00067>
- O'Connor, S. & ChatGPT. (2023) *Educación en enfermería. Practica* 66, 103537
- O'Mahoney, K. & Heinz, M. (2016). Using comment only marking in a mathematics classroom in the Republic of Ireland: experience and learning of a student teacher. *Journal of Teacher Action Research*, 2(2), 58-69.
- Pedreño. A. (2023). *ChatGPT y las universidades*. <https://www.universidadsi.es/chatgpt-y-las-universidades/>

- Petty, G. (2009). *Evidence Based Teaching – a practical approach* (2nd edition). Cheltenham: Nelson Thornes. [http://155.0.32.9:8080/jspui/bitstream/123456789/91/1/%5BGeoff\\_Petty%5D\\_Evidence-based\\_Teaching\\_A\\_Practical%28BookFi.org%29.pdf](http://155.0.32.9:8080/jspui/bitstream/123456789/91/1/%5BGeoff_Petty%5D_Evidence-based_Teaching_A_Practical%28BookFi.org%29.pdf)
- Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Mejora diferencial en la comprensión de los estudiantes de los principios matemáticos después de una intervención de evaluación formativa. *The Journal of Educational Research*, 104 (5), 330-339.
- Race, P. (2018). Is the ‘time of the assessed essay’ over? *University of Sussex blog article*. <https://blogs.sussex.ac.uk/business-school-teaching/2018/11/14/is-the-time-of-the-assessed-essay-over/>
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science* 28(1), 4-13.
- Reigeluth, C. M., Ed. (1999). *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*. Mahwah, NJ, Lawrence Erlbaum Associates Publishers.
- Reigeluth, C. (2012). Teoría instruccional y tecnología para el nuevo paradigma de la educación. *RED, Revista de Educación a Distancia*. 32. <http://www.um.es/ead/red/32>
- Saul, D. (2023). ‘AI Is The New Electricity’: Bank Of America Picks 20 Stocks To Cash In On ChatGPT Hype. *Forbes BUSINESS*. <https://www.forbes.com/sites/dereksaul/2023/03/01/ai-is-the-new-electricity-bank-of-america-picks-20-stocks-to-cash-in-on-chatgpt-hype/>
- Science, Technology, and Public Policy. (2022). Parthasarathy discusses implications of Large Language Models. <https://stpp.fordschool.umich.edu/news/2022/parthasarathy-discusses-implications-large-language-models>
- Sharples, M. (2022) Automated Essay Writing: An AIED Opinion. *Int J Artif Intell Educ.* 32, 1119–1126. <https://doi.org/10.1007/s40593-022-00300-7>
- Sharples. (2023). Twitter <https://twitter.com/sharplm/status/1615607290320519169>
- Shepard, L. A. (2005). Formative assessment: Caveat emptor. In *ETS Invitational Conference The Future of Assessment: Shaping Teaching and Learning*, New York. [https://web.archive.org/web/20111007220347/http://www.cpre.org/ccii/images/stories/ccii\\_pdfs/shepard%20formative%20assessment%20caveat%20emptor.pdf](https://web.archive.org/web/20111007220347/http://www.cpre.org/ccii/images/stories/ccii_pdfs/shepard%20formative%20assessment%20caveat%20emptor.pdf)
- Siemens, G. (2014, May). Multiple pathways: Blending xMOOCs & cMOOCs. <http://www.elearnspace.org/blog/2014/05/06/multiple-pathways-blending-xmoocscmoocs/>
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—Should professors worry?[published online ahead of print December 9, 2022]. *Nature News*.
- Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, 613(7945): 620-621. <https://doi.org/10.1038/d41586-023-00107-z>
- Surden, H. (2019). "Inteligencia artificial y derecho: una descripción general", *Revista de derecho de la Universidad Estatal de Georgia*, vol. 35. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3411869#](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3411869#).
- Tennyson, R. D., Schott, F., Dijkstra, S., & Seel, N. M. (Eds.). (1997). *Instructional Design: International Perspectives. Theory, research, and models. Vol. 1* (Vol. 1). Routledge.
- Thompson, M., & Wiliam, D. (2007). Tight but loose: A conceptual framework for scaling up school reforms. In *annual meeting of the American Educational Research Association*. Chicago, IL. <https://cpb-ap-se2.wpmucdn.com/global2.vic.edu.au/dist/d/30223/files/2017/01/Tight-but-Loose-AERA-07-2di57v1.pdf>
- Turnitin (2020). How Teachers Can Prepare for AI-Based Writing. *Turnitin blog article*. <https://www.turnitin.com/blog/how-teachers-can-prepare-for-ai-based-writing>
- Weller, M. (2013). Good post by Martin Weller [@mweller] on Coursera's new plans - The Ed Techie: *You can stop worrying about MOOCs now.*

- [http://nogoodreason.typepad.co.uk/no\\_good\\_reason/2013/05/you-can-stop-worrying-about-moocs-now.html](http://nogoodreason.typepad.co.uk/no_good_reason/2013/05/you-can-stop-worrying-about-moocs-now.html)
- Wiliam, D. (2000). *Integrating summative and formative functions of assessment*. [https://discovery.ucl.ac.uk/id/eprint/1507176/1/Wiliam2000IntergratingAEA-E\\_2000\\_keynoteaddress.pdf](https://discovery.ucl.ac.uk/id/eprint/1507176/1/Wiliam2000IntergratingAEA-E_2000_keynoteaddress.pdf)
- Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. *Nature*, 1, 20.
- Wing, J. M. (2006). Computational thinking. it represents a universally applicable attitude and skill set everyone, not just computer scientists, would be eager to learn and use. *Commun. ACM* 49(3). <https://doi.org/10.1109/vlhcc.2011.6070404>
- Wing, J. M. (2008). Computational thinking and thinking about computing. *The Royal Society Publishing*. <https://doi.org/10.1098/rsta.2008.0118>
- Zapata-Ros, M. (2012). Calidad en entornos ubicuos de aprendizaje. *Revista de Educación a Distancia (RED)*, (31). <https://revistas.um.es/red/article/view/232871>
- Zapata-Ros, M. (2012). *Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Bases para un nuevo modelo teórico a partir de una visión crítica del “conectivismo”*. <http://eprints.rclis.org/17463/>
- Zapata-Ros, M. (2013). *Charles Reigeluth: la personalización del aprendizaje y el nuevo paradigma de la educación para la sociedad postindustrial del conocimiento.*, <http://eprints.rclis.org/19406/>
- Zapata-Ros, M. (2013). Enseñanza Universitaria en línea, MOOC y aprendizaje divergente. *Academia.edu* (preprint) [https://www.academia.edu/6190225/Ense%C3%B1anza\\_Universitaria\\_en\\_l%C3%ADnea\\_MOOC\\_aprendizaje\\_divergente\\_y\\_creatividad\\_Online\\_Higher\\_Education\\_MOOC\\_divergent\\_learning\\_and\\_creativity](https://www.academia.edu/6190225/Ense%C3%B1anza_Universitaria_en_l%C3%ADnea_MOOC_aprendizaje_divergente_y_creatividad_Online_Higher_Education_MOOC_divergent_learning_and_creativity)
- Zapata-Ros, M. (2013). La evaluación en el nuevo paradigma de la educación en la Sociedad Postindustrial del Conocimiento. <http://eprints.rclis.org/20157/>
- Zapata-Ros, M. (2013b). El “problema de 2 sigma” y el aprendizaje ayudado por la tecnología. *E-Lis Repository*. <http://eprints.rclis.org/18866/8/MOOCs.pdf>
- Zapata-Ros, M. (2013c). Las teorías del aprendizaje y el diseño instruccional. El esquema incompleto. <http://redesabiertas.blogspot.com.es/2013/04/las-teorias-del-aprendizaje-y-el-diseno.html>
- Zapata-Ros, M. (2015). El diseño instruccional de los MOOC y el de los nuevos cursos abiertos personalizados. *Revista de Educación a Distancia*, (45). <https://revistas.um.es/red/article/download/238661/181351>
- Zapata-Ros, M. (2018a). El modelo educativo de Singapur y el «mastery learning» de Bloom. *Blog académico RED. El aprendizaje en la sociedad del conocimiento*. <https://red.hypotheses.org/1261>
- Zapata-Ros, M. (2018b). La universidad inteligente. *RED. Revista de Educación a Distancia*, 57(10). <http://www.um.es/ead/red/57/zapata2.pdf> <http://dx.doi.org/10.6018/red/57/10>
- Zapata-Ros, M. (2020). ¿Qué tienen de común Jeff Bezos y Benjamin Bloom? / What do Jeff Bezos and Benjamin Bloom have in common? 10.13140/RG.2.2.14749.26085
- Zapata-Ros, M. (2022). *El pensamiento computacional, la competencia clave en una nueva alfabetización*. New York: Create Space Independent Publishing.
- Zapata-Ros, M., & Palacios, Y. B. (2021). El pensamiento bayesiano, un pensamiento computacional omnipresente. *Revista de Educación a Distancia (RED)*, 21(68).