

**Ciencia Latina**  
Internacional

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.  
ISSN 2707-2207 / ISSN 2707-2215 (en línea), septiembre-octubre 2024,  
Volumen 8, Número 5.

[https://doi.org/10.37811/cl\\_rcm.v8i5](https://doi.org/10.37811/cl_rcm.v8i5)

**RENDIMIENTO ACADÉMICO EN UN  
PROGRAMA DE ENSEÑANZA DE JAVA  
UTILIZANDO SOFTWARE TUTOR BASADO  
EN CHATGPT: ESTUDIO DE CASO**

ACADEMIC PERFORMANCE IN TEACHING  
JAVA PROGRAMMING LANGUAGE USING A CHATGPT-  
BASED TUTOR SOFTWARE: A CASE STUDY

**Félix Fernández-Peña**

Universidad Técnica de Ambato, Ecuador

**Marlena León-Mendoza**

Universidad de las Américas, Ecuador

DOI: [https://doi.org/10.37811/cl\\_rcm.v8i5.14595](https://doi.org/10.37811/cl_rcm.v8i5.14595)

## Rendimiento Académico en un Programa de Enseñanza de Java Utilizando Software Tutor Basado en ChatGPT: Estudio de Caso

**Félix Fernández Peña<sup>1</sup>**

[fo.fernandez@uta.edu.ec](mailto:fo.fernandez@uta.edu.ec)

<https://orcid.org/0000-0003-0834-3377>

Universidad Técnica de Ambato  
Ecuador

**Marlena León Mendoza**

[marlena.leon@udla.edu.ec](mailto:marlena.leon@udla.edu.ec)

<https://orcid.org/0000-0003-4944-9709>

Universidad de las Américas  
Ecuador

### RESUMEN

La enseñanza de lenguajes de computación es un reto que enfrentan los estudiantes de ingeniería de software a nivel mundial. En este proceso, contar con la ayuda de un tutor que responda preguntas referentes a cómo programar ciertas funcionalidades puede ser fundamental para desarrollar las capacidades de abstracción requeridas. Con este trabajo presentamos el resultado de una investigación que ha pretendido evaluar el impacto del uso de un software tutor basado en ChatGPT para aprender lenguaje Java. El estudio se llevó a cabo con la participación de 40 estudiantes que se dividieron en dos grupos de 20. El rendimiento académico se analizó utilizando un estadígrafo t-student. El resultado permitió comprobar que la diferencia en el rendimiento académico de ambos grupos fue significativa con un valor  $p < 0.05$ , y que el uso del software tutor mejoró significativamente el proceso de aprendizaje de los estudiantes.

**Palabras clave:** software tutor, IA generativa, rendimiento académico

---

<sup>1</sup> Autor principal.

Correspondencia: [fo.fernandez@uta.edu.ec](mailto:fo.fernandez@uta.edu.ec)

# Academic Performance in teaching Java programming language using a ChatGPT-based tutor software: a case study

## ABSTRACT

Teaching programming languages is a challenge for the students of software engineering around the world. Having the help of a tutor who answers questions about how to program certain functionalities can be essential for developing the required abstraction skills. We present the results in this research focused on measuring the impact of using a ChatGPT-based software for tutoring the students while learning how to program in Java. The study was carried out with the participation of 40 students which were randomly assigned to one of two groups of 20 students each. Academic performance at the end of the course was analysed using a t-student test. The result allowed to probed the hyphotesis of the research. The academic performance is significantly dependent on the use of the ChatGPT-based software for tutoring with a p-value of 0.05.

**Keywords:** tutor, generative AI, academic performance

*Artículo recibido 10 septiembre 2024*

*Aceptado para publicación: 12 octubre 2024*



## INTRODUCCIÓN

Las habilidades, tanto con lenguajes de programación como con idiomas, son fundamentales para estudiantes de la carrera de Ingeniería de Software (Borges & de Souza, 2024) (Yeo et al., 2024). Por su parte, las herramientas de IA generativa, como ChatGPT, son capaces tanto de generar un discurso coherente en un idioma dado, como de programar en múltiples lenguajes de programación de manera satisfactoria. En este sentido, Fernández-Peña et al. (en prensa) manifiestan que existe una correlación positiva significativa entre el nivel de exposición a las IA generativas y su uso real en el ámbito educativo. Este último resultado indica que, con el actual interés que genera el uso de las IA generativas (Albayati, 2024), el futuro será de un innegable uso de este tipo de tecnologías.

Si bien no existe un consenso en cuanto a la importancia que tienen las habilidades en matemáticas y/o en idiomas durante el proceso de aprendizaje (Malt, 2024), lo cierto es que todo lenguaje, sea humano o de programación, constituye un sistema estructurado de símbolos y reglas para la comunicación o la codificación de información (Malt, 2024) (Yeo et al., 2024). Tomando esta generalización como premisa, dado el éxito actual de las IA generativas en el apoyo al proceso de aprendizaje (Albayati, 2024) (Alberth, 2023) (Rejeb et al., 2024) (Pérez-Imaicela et al., en prensa) (Martínez-Olmo & González, 2024) (Sun, 2024) (Shahzad, Xu & Javed, 2024), el presente trabajo tiene como objetivo evaluar el impacto que este tipo de herramientas tuvo en la enseñanza del lenguaje de programación Java en un curso durante el semestre octubre 2023 a febrero 2024.

El estudio llevado a cabo se sustentó en el uso de la herramienta tutor propuesta por Pérez-Imaicela, Coello-Fiallos & Fernández-Peña (en prensa), la misma que sirve como intermediario entre el estudiante y ChatGPT, como IA generativa escogida para llevar a cabo este estudio. Esta herramienta permitió corroborar el uso que los estudiantes le dieron a la IA generativa durante el proceso de aprendizaje y retroalimentar al docente con relación a la forma en que lo vinieron haciendo durante el semestre de clases.

El presente trabajo complementa los resultados de Pérez-Imaicela, Coello-Fiallos & Fernández-Peña (en prensa), por cuanto su trabajo se centró en la evaluación del nivel de usabilidad de la herramienta y no en el rendimiento académico alcanzado por los estudiantes.



Esta evaluación es pertinente por cuanto no encontramos en la literatura una evaluación del rendimiento académico logrado en cursos de lenguajes de programación en que se haya utilizado IA generativas.

Para el desarrollo de la investigación se planteó como hipótesis la siguiente:

**Hipótesis Nula (H0):** No hay diferencia significativa en el rendimiento académico promedio entre los estudiantes que utilizaron un software tutor basado en ChatGPT y los que no.

**Hipótesis Alternativa (H1):** Hay una diferencia significativa en el rendimiento académico promedio entre los estudiantes que utilizaron un software tutor basado en ChatGPT y los que no.

## **METODOLOGÍA**

La investigación del presente trabajo se dividió en dos fases. En una primera fase se evaluó la calidad de las respuestas del software tutor implementado con ChatGPT. En una segunda fase se intentó comprobar la hipótesis de investigación. El enfoque de investigación seguido ha sido cuantitativo experimental. Se trabajó con un grupo de 40 estudiantes a los que se les dividió aleatoriamente en dos grupos de 20 estudiantes cada uno. El grupo experimental utilizó el software tutor desarrollado por Pérez et al. (en prensa), mientras que el grupo control no lo utilizó, continuando con el método de estudio de lenguaje Java tradicional con apoyo del docente en tutorías en horarios preestablecidos. En la primera fase se cuantificó, con criterio del docente, como experto en la temática, la calidad de las respuestas de ChatGPT. En la segunda fase se utilizó la prueba T-Student para determinar si la diferencia en las medias de rendimiento entre ambos grupos fue estadísticamente significativa.

La calificación final obtenida (CF) en el curso fue lo que se utilizó como indicador de rendimiento académico. Este valor tuvo en cuenta el resultado en actividades prácticas individuales (PI), actividades prácticas en grupo (PG) y evaluaciones teóricas (ET), de acuerdo a la fórmula de cálculo 1, que se muestra a continuación:

$$CF = 0.3 * PI + 0.3 * PG + 0.4 * ET \quad (1)$$

Evidentemente la enseñanza de un lenguaje de programación le da mayor peso al componente de aprendizaje basado en actividades. Coincidente con el resultado de Al Shloul et al. (2024), es en este ámbito que se espera que la herramienta ChatGPT tenga un mayor impacto en el rendimiento académico.



En este sentido, tenemos en cuenta las limitaciones identificadas por Bucaioni et al. (2024) para contextos de mayor complejidad pero que no limitan la utilidad en cursos introductorios y medios de programación.

## RESULTADOS Y DISCUSIÓN

Los estudiantes del grupo experimental contaron con un tutor en línea disponible las 24 horas del día, capaz de responder preguntas del curso de programación en cuestión. En este sentido, evidentemente la atención diferenciada a cada estudiante en temas técnicos fue mayor en el grupo experimental que en el grupo de control.

En cuando a la calidad de las respuestas del software tutor, se llevó a cabo un muestreo aleatorio de 30 respuestas ofrecidas por ChatGPT a preguntas referentes a sintaxis y generación de sentencias simples, comprobando que la exactitud de las respuestas fue del 100%. En la figura 1 se muestra el resultado obtenido al pedirle a ChatGPT la generación de una sentencia con el siguiente prompt: *Dado que tengo una lista simplemente enlazada cuyo primer nodo es "primero", cuál es la sentencia que me permite declarar un nodo "aux" que haga referencia al tercer nodo de dicha lista?*. En este caso, el estudiante solicitó una sentencia y la herramienta genera una respuesta que tiene en cuenta las condiciones que pueden producirse para invalidar dicha sentencia. Es decir que el resultado no se limitó a lo solicitado sino a lo que requerido para que funcione de manera correcta. En este sentido, la tutoría de ChatGPT en el proceso de aprendizaje del estudiante fue excelente. Este resultado indica un nivel de calidad adecuado para aprender a programar estructuras de datos en Java.

**Figura 1.** Respuesta de ChatGPT en la generación de una sentencia.

```
En un lenguaje de programación como Java, por ejemplo, si tienes una clase Nodo que representa un nodo de la lista simplemente enlazada con un atributo siguiente que apunta al siguiente nodo, podrías hacer algo similar al siguiente código:
```

```
Nodo aux = primero; // Inicializamos auxiliar con el primer nodo
if (aux != null) {
    aux = aux.getSiguiente(); // Avanzamos al segundo nodo
    if (aux != null) {
        aux = aux.getSiguiente(); // Avanzamos al tercer nodo
    }
}
```

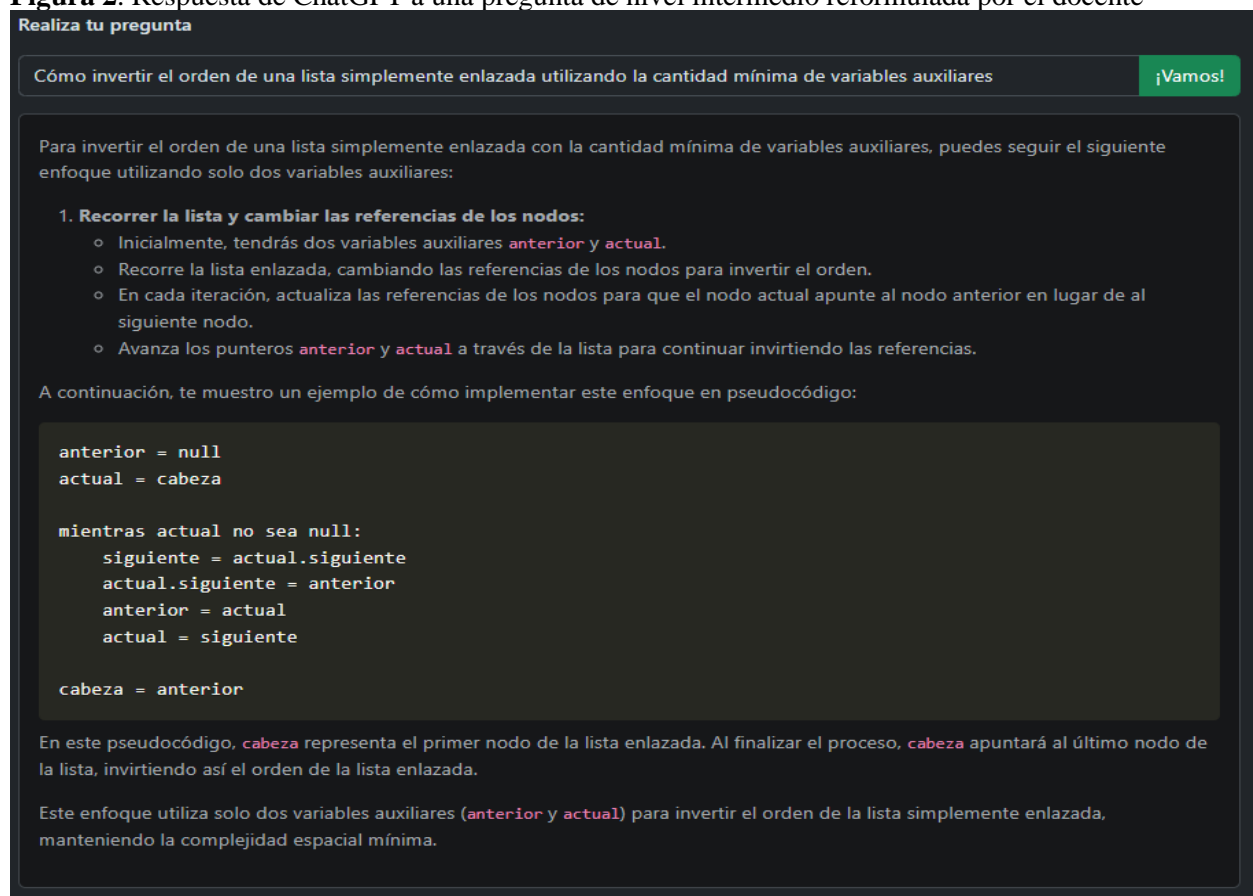
Del mismo modo, se revisó la respuesta que dio la IA generativa a casos de estudio de mediana dificultad. En este caso, del total de 30 respuestas escogidas aleatoriamente, se pudo comprobar que 9 de ellas no fueron satisfactorias. Para este 30% de casos, se reformuló la pregunta del estudiante y se

logró un resultado satisfactorio por parte de la IA generativa (la figura 2 muestra el resultado de ChatGPT a una de las preguntas reformuladas por el docente).

Este resultado hace pensar que para lograr un buen resultado al utilizar la IA generativa en la solución de casos de mediana complejidad se requiere de un mayor conocimiento de parte del estudiante. Dicho de otra manera, al aumentar la complejidad de la solución, el uso de la IA requiere de una mayor habilidad en la formulación del *prompt* por parte del usuario. En este sentido, Yu (2024) sugiere el cambio del rol del docente, del que posiblemente se demande cada vez más su atención al diseño de prompts de acuerdo a la temática que enseña.

El trabajo de Chiu (2024) también se centra en un conjunto de recomendaciones para la transformación de la educación superior utilizando las IA generativas que demanda el cambio del rol del docente. Así mismo, Zhai, Wibowo & Li (2024), Essel et al. (2024) y De Santis, Martino & Rizzi (2024) destacan la importancia del fortalecimiento de las habilidades cognitivas para la interacción con este tipo de tecnologías.

**Figura 2.** Respuesta de ChatGPT a una pregunta de nivel intermedio reformulada por el docente



Realiza tu pregunta

Cómo invertir el orden de una lista simplemente enlazada utilizando la cantidad mínima de variables auxiliares ¡Vamos!

Para invertir el orden de una lista simplemente enlazada con la cantidad mínima de variables auxiliares, puedes seguir el siguiente enfoque utilizando solo dos variables auxiliares:

- Recorrer la lista y cambiar las referencias de los nodos:**
  - Inicialmente, tendrás dos variables auxiliares **anterior** y **actual**.
  - Recorre la lista enlazada, cambiando las referencias de los nodos para invertir el orden.
  - En cada iteración, actualiza las referencias de los nodos para que el nodo actual apunte al nodo anterior en lugar de al siguiente nodo.
  - Avanza los punteros **anterior** y **actual** a través de la lista para continuar invirtiendo las referencias.

A continuación, te muestro un ejemplo de cómo implementar este enfoque en pseudocódigo:

```
anterior = null
actual = cabeza

mientras actual no sea null:
    siguiente = actual.siguiente
    actual.siguiente = anterior
    anterior = actual
    actual = siguiente

cabeza = anterior
```

En este pseudocódigo, **cabeza** representa el primer nodo de la lista enlazada. Al finalizar el proceso, **cabeza** apuntará al último nodo de la lista, invirtiendo así el orden de la lista enlazada.

Este enfoque utiliza solo dos variables auxiliares (**anterior** y **actual**) para invertir el orden de la lista simplemente enlazada, manteniendo la complejidad espacial mínima.

En cuanto al análisis para comprobar la hipótesis del presente estudio, la tabla 1 muestra los resultados del procesamiento estadístico descriptivo de la calificación final de los estudiantes en el curso. Con un nivel de significancia ( $\alpha$ ) de 0.05, y un valor de grados de libertad ( $df$ ) de 38, dado que son dos grupos, de 20 estudiantes cada uno, teniendo en cuenta la fórmula de cálculo 2. Este resultado, consultado en la tabla t bilateral.

$$df = n1 + n2 - 2 \quad (2)$$

**Tabla 1.** Estadística descriptiva de los resultados académicos de los estudiantes.

Grupo	$\bar{X}$	$s$	$n$
Experimental ( $x_1$ )	7.99	0.137477	20
De control ( $x_2$ )	7.3	0.371484	20

Sustituyendo los valores en la fórmula correspondiente al cálculo del estadígrafo (fórmula 3), se obtuvo un valor  $t \approx 7.82$ .

$$t = \frac{X_1 - X_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (3)$$

Considerando un nivel de significancia de 0.05 (5%) y grados de libertad  $df=n_1+n_2-2=38$ , el valor crítico  $t$  para una prueba bilateral es aproximadamente 2.024. Como el valor calculado de  $t=7.82$  es mayor que el valor crítico de 2.024, se rechaza la hipótesis nula. La prueba T-Student revela una diferencia significativa entre las medias de ambos grupos, con un valor de  $t$  de 7.82 ( $p < 0.05$ ). Esto indica que las medias de los grupos son significativamente diferentes, sugiriendo que el resultado académico de los grupos sí puede estar relacionado con el uso o no de la herramienta ChatGPT como tutor de programación Java.

## CONCLUSIONES

Como resultado del presente trabajo se ha logrado comprobar una diferencia significativa en el rendimiento académico promedio entre los estudiantes que utilizaron un software tutor basado en ChatGPT y los que no. Así mismo, se evaluaron un total de 60 respuestas de ChatGPT, comprobando, basados en el criterio del docente, que las respuestas de la IA generativa fueron satisfactorias al funcionar como tutor de programación en lenguaje Java en un nivel básico.





Para un nivel más avanzado, las respuestas son igual de correctas cuando se tiene un nivel de conocimiento que permita tener en cuenta aspectos específicos necesarios al formular el prompt a la IA. No obstante, los resultados de esta investigación no son concluyentes. Se sugiere llevar a cabo un estudio transversal que involucre mayor cantidad de estudiantes, y un estudio longitudinal para validar la utilidad o no de este tipo de herramientas en el aprendizaje de lenguajes de programación.

## REFERENCIAS BIBLIOGRAFICAS

- Al Shloul, T., Mazhar, T., Abbas, Q., Iqbal, M., Ghadi, Y. Y., Shahzad, T., ... Hamam, H. (2024). Role of activity-based learning and ChatGPT on students' performance in education. *Computers and Education: Artificial Intelligence*, 6, 100219.  
<https://doi.org/https://doi.org/10.1016/j.caeai.2024.100219>
- Albayati, H. (2024). Investigating undergraduate students' perceptions and awareness of using ChatGPT as a regular assistance tool: A user acceptance perspective study. *COMPUTERS AND EDUCATION: ARTIFICIAL INTELLIGENCE*, 100203. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100203>
- Alberth. (2023). The use of ChatGPT in academic writing: a blessing or a curse in disguise? *TEFLIN JOURNAL*, 34(2), 337–352. <https://doi.org/10.15639/teflinjournal.v34i2/337-352>
- Borges, G. G., & de Souza, R. C. G. (2024). Skills development for software engineers: Systematic literature review. *INFORMATION AND SOFTWARE TECHNOLOGY*, 168. <https://doi.org/10.1016/j.infsof.2023.107395>
- Bucaioni, A., Ekedahl, H., Helander, V., & Nguyen, P. T. (2024). Programming with *ChatGPT*: How far can we go? *MACHINE LEARNING WITH APPLICATIONS*, 100526. <https://doi.org/https://doi.org/10.1016/j.mlwa.2024.100526>
- Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *COMPUTERS AND EDUCATION: ARTIFICIAL INTELLIGENCE*, 6. <https://doi.org/10.1016/j.caeai.2023.100197>
- De Santis, E., Martino, A., & Rizzi, A. (2024). Human Versus Machine Intelligence: Assessing Natural Language Generation Models Through Complex Systems Theory. *IEEE TRANSACTIONS ON*



*PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 46(7), 4812–4829.

<https://doi.org/10.1109/TPAMI.2024.3358168>

- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *COMPUTERS AND EDUCATION: ARTIFICIAL INTELLIGENCE*, 6, 100198. <https://doi.org/https://doi.org/10.1016/j.caeai.2023.100198>
- Fernández Peña, F., Moreno-Guamán, Y., Urrutia-Urrutia, P., Tigse-Bravo, W., León-Mendoza, M. (en prensa). Academic Essays and the Use of Generative AI in Social Science Subjects. *8TH INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY & SYSTEMS*, 2025.
- Malt, B. C. (2024). Representing the World in Language and Thought. *TOPICS IN COGNITIVE SCIENCE*, 16(1), 6–24. <https://doi.org/10.1111/tops.12719>.
- Martínez-Olmo, F., González Catalán, F. (2024). Systematic review of trends in the application of artificial intelligence to the field of academic writing in the social sciences. *DIGITAL EDUCATION REVIEW*, Junio – 45. <https://doi.org/10.1344/der.2024.45.37-42>.
- Pérez-Imaicela, R., Coello-Fiallos, D., & Fernández-Peña, F. (en prensa). Academic reinforcement software tutor based on ChatGPT. An approach evaluated for software programming courses. *MULTIDISCIPLINARY INTERNATIONAL CONFERENCE ON RESEARCH APPLIED TO DEFENSE AND SECURITY*, 2024.
- Rejeb, A., Rejeb, K., Appolloni, A., Treiblmaier, H., & Iranmanesh, M. (2024). Exploring the impact of ChatGPT on education: A web mining and machine learning approach. *THE INTERNATIONAL JOURNAL OF MANAGEMENT EDUCATION*, 22(1), 100932. <https://doi.org/https://doi.org/10.1016/j.ijme.2024.100932>
- Shahzad, M.F., Xu, S. & Javed, I. ChatGPT awareness, acceptance, and adoption in higher education: the role of trust as a cornerstone. *INTERNATIONAL JOURNAL OF EDUCATIONAL TECHNOLOGY IN HIGHER EDUCATION*, 21 (46) (2024). <https://doi.org/10.1186/s41239-024-00478-x>



- Sun, D., Looi, C.-K., Li, Y., Zhu, C., Zhu, C., & Cheng, M. (2024). Block-based versus text-based programming: a comparison of learners' programming behaviors, computational thinking skills and attitudes toward programming. *ETR&D-EDUCATIONAL TECHNOLOGY RESEARCH AND DEVELOPMENT*, 72(2), 1067–1089. <https://doi.org/10.1007/s11423-023-10328-8>.
- Yeo, S., Ma, Y.-S., Kim, S. C., Jun, H., & Kim, T. (2024). Framework for evaluating code generation ability of large language models. *ETRI JOURNAL*, 46(1), 106–117. <https://doi.org/10.4218/etrij.2023-0357>.
- Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: New demands for teachers' roles. *HELIYON*, e24289. <https://doi.org/10.1016/j.heliyon.2024.e24289>
- Zhai, C., Wibowo, S. & Li, L.D. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *SMART LEARNING ENVIRONMENTS*, 11 (28) (2024). <https://doi.org/10.1186/s40561-024-00316-7>

