



Efectos de un videojuego en la mejora del pensamiento computacional de estudiantes universitarios

- Effects of a Video Game on Enhancing Computational Thinking in University Students
- Efeitos de um videogame na melhoria do pensamento computacional de estudantes universitários

Forma de citar este artículo:

Moncada-Beltrán, C. D., Mariaca-Orozco, O. O. y Rivera-Pinzón, D. M. (2025). Efectos de un videojuego en la mejora del pensamiento computacional de estudiantes universitarios. *Tecné, Episteme y Didaxis*: TED, (58), 51 - 68. <https://doi.org/10.17227/ted.num58-22611>

Resumen

Este estudio evaluó el efecto de un videojuego de realidad aumentada en el desarrollo del pensamiento computacional en estudiantes de primer semestre de Ingeniería en la Universidad Piloto de Colombia. Participaron 38 estudiantes, seleccionados de una población de 41, lo que garantiza un 95% de confianza y un margen de error del 5%. Inicialmente, se nivelaron los conocimientos matemáticos básicos y se aplicó un primer test con preguntas del examen SABER 11 para medir el pensamiento computacional, registrando tiempos y efectividad de respuestas. Posteriormente, los estudiantes utilizaron una aplicación de videojuego de realidad aumentada, diseñada con UNITY, que jugaron 10 veces. Finalmente, se aplicó un segundo test similar al primero, con preguntas reorganizadas y uso de sinónimos. Los resultados mostraron una disminución significativa en el tiempo promedio de respuesta entre los test de 13,2 minutos ($p < 0.05$) y un aumento del 15% en la efectividad de las respuestas. El coeficiente KR20 obtenido fue de 0.62, indicando una consistencia interna aceptable. Estos hallazgos sugieren que el uso de videojuegos de realidad aumentada puede ser una herramienta efectiva para mejorar el pensamiento computacional en estudiantes universitarios.

Palabras clave

resolución de problemas; educación matemática; aplicaciones informáticas; operaciones formales; software informático

Carlos David Moncada-Beltrán*  
Oscar Orlando Mariaca-Orozco**  
Diego Mauricio Rivera-Pinzón**  

* Especialista en Desarrollo Organizacional, estudiante de la Maestría en Tecnologías de la Información Aplicadas a la Educación (MTIAE), Universidad Pedagógica Nacional, Bogotá, Colombia. cdmoncadab@upn.edu.co

** Especialista en Teleinformática, estudiante Maestría en Tecnologías de la Información Aplicadas a la Educación (MTIAE), Universidad Pedagógica Nacional, Bogotá, Colombia. oomariacao@upn.edu.co

*** Magíster en Ingeniería - Automatización Industrial. Universidad Pedagógica Nacional. Departamento de tecnología. Bogotá, Colombia. drrivera@pedagogica.edu.co

Artículo de Investigación

Fecha de recepción: 24/12/2024
Fecha de aprobación: 10/04/2025
Fecha de publicación: 01/07/2025



Abstract

This study evaluated the effect of an augmented reality video game on the development of computational thinking in first-semester engineering students at Universidad Piloto de Colombia. A total of 38 students participated, selected from a population of 41, ensuring a 95% confidence level and a 5% margin of error. Initially, students' basic mathematical knowledge was standardised, and a pre-test using questions from the SABER 11 exam was administered to assess computational thinking, recording response times and accuracy. Subsequently, the students engaged with an augmented reality video game application developed using UNITY, which they played ten times. A second test, similar to the first but with restructured questions and synonyms, was then administered. The results showed a significant reduction in the average response time between the tests—by 13.2 minutes ($p < 0.05$)—and a 15% increase in response accuracy. The KR20 coefficient obtained was 0.62, indicating acceptable internal consistency. These findings suggest that augmented reality video games may be an effective tool for enhancing computational thinking in university students.

Keywords

problem solving; mathematics education; computer applications; formal operations; computer software

Resumo

Este estudo avaliou o efeito de um videogame de realidade aumentada no desenvolvimento do pensamento computacional em estudantes do primeiro semestre do curso de engenharia da Universidade Piloto da Colômbia. Participaram 38 estudantes, selecionados de uma população de 41, garantindo um nível de confiança de 95% e uma margem de erro de 5%. Inicialmente, nivelaram-se os conhecimentos matemáticos básicos e aplicou-se um teste com perguntas do exame SABER 11 para medir o pensamento computacional, registrando-se os tempos e a efetividade das respostas. Em seguida, os estudantes utilizaram um aplicativo de videogame de realidade aumentada, desenvolvido com o UNITY, que jogaram dez vezes. Por fim, foi aplicado um segundo teste, semelhante ao primeiro, mas com questões reorganizadas e uso de sinônimos. Os resultados mostraram uma redução significativa no tempo médio de resposta entre os testes—13,2 minutos ($p < 0,05$)—e um aumento de 15% na efetividade das respostas. O coeficiente KR20 obtido foi de 0,62, indicando consistência interna aceitável. Esses achados sugerem que o uso de videogames de realidade aumentada pode ser uma ferramenta eficaz para melhorar o pensamento computacional de estudantes universitários.

Palavras-chave

resolução de problemas; ensino de matemática; aplicações computacionais; operações formais; software computacional

Introducción

El pensamiento computacional (PC) es una habilidad para descomponer problemas complejos, reconocer patrones, abstraer y generar algoritmos. Claves para el desarrollo de competencias en la resolución de problemas de razonamiento, especialmente en entornos educativos universitarios donde la tecnología y el análisis lógico desempeñan un papel central. En este contexto, la presente investigación busca determinar si un entrenamiento específico en PC, mediante el uso de videojuegos educativos, puede mejorar significativamente el desempeño de los estudiantes en pruebas certificadas, específicamente en el área de matemáticas.

El estudio se llevó a cabo con una muestra de 38 estudiantes que participaron en un experimento controlado, dividido en grupos de intervención y control, donde se evaluó su capacidad para resolver problemas matemáticos antes y después de jugar con el aplicativo. Se diseñaron dos pruebas basadas en preguntas del examen SABER 11, homologado internacionalmente, que mide competencias clave, incluido el PC.

Además, se aplicaron herramientas estadísticas como la prueba t de Student para muestras dependientes y el coeficiente de Kuder-Richardson para evaluar la consistencia interna de las pruebas. Los resultados preliminares muestran una disminución significativa en los tiempos de respuesta y un incremento notable en tiempo de respuesta de manera correcta. Esta investigación no solo explora la efectividad del PC como herramienta educativa, sino también el impacto del uso de videojuegos en el aprendizaje formal.

Antecedentes

En las plataformas web se pueden encontrar diversas propuestas relacionadas con VJ y

actividades orientadas al Pensamiento Computacional (PC) y la programación gráfica. Por ejemplo, Araujo *et al.* (2019) presentan una plataforma web llamada MICAS (<https://micas.epl.di.uminho.pt/>), diseñada para promover estas habilidades. Por otro lado, Villacís (2014) desarrolla una propuesta enfocada en actividades de videojuego (VJ) dirigidas a estudiantes de K12 (desde nivel preescolar hasta el grado 12), en materias como matemáticas, tecnología e informática. Los Videojuegos son aplicaciones software altamente interactivas en las que una experiencia lúdica se lleva a cabo en un dispositivo informático (Fabregat, 2013). Así mismo, se destacan proyectos que utilizan entornos gráficos como Unity 3D junto con realidad aumentada (AR) y marcadores, una tendencia en el desarrollo de software educativo, tal como se observa en el trabajo de Alkinformática. La realidad aumentada es un sistema que permite la visualización de una capa virtual superpuesta sobre elementos del mundo real (Azuma, 1997). Lin y Chen (2020) resaltan la importancia de la programación como una herramienta para despertar e inspirar el potencial de los estudiantes. Por su parte, Muñoz *et al.* (2018) proponen actividades didácticas que fomentan el pensamiento computacional en niños y adolescentes, con un enfoque especial en estudiantes con autismo. Csernoch *et al.* (2021) presentan un artículo detallando un proyecto de cuatro años que evalúa la eficacia de la enseñanza mediante el uso de hojas de cálculo denominadas Sprego. Estas herramientas, diseñadas con un enfoque en la resolución de problemas conceptuales, demostraron ser más efectivas que los métodos tradicionales, ofreciendo conocimientos más sólidos y duraderos a los estudiantes.

Medir el nivel de desarrollo del PC de un estudiante debe hacerse de forma indirecta ya que es una variable latente y, como tal, no es medible de forma directa según lo clasifica

Datalab (2003). El trabajo de Mendoza *et al.* (2023) creó un diagnóstico del PC como una herramienta que puede identificar a estudiantes de 1er año de ingeniería en riesgo. El estudio se realizó con la participación de 469 estudiantes de tres instituciones de Estados Unidos, todos ellos realizaron el diagnóstico. Los análisis factoriales confirmatorios dieron como resultado que 20 ítems generan un buen rango de ajuste del modelo, con un coeficiente de confiabilidad de consistencia interna, α de Cronbach de 0,86. Chai *et al.* (2019) proponen una nueva forma de evaluación llamada: Sistema dinámico de evaluación ponderada: DWES, por sus siglas en inglés. La propuesta consta, primero, de definir un nuevo criterio de evaluación del PC abordando 8 aspectos. En segundo lugar, con base en los criterios propuestos, se presentó una herramienta de análisis que automáticamente evalúa las habilidades del PC de los proyectos de los alumnos. En tercer lugar, teniendo en cuenta las características de los proyectos, se ajustaron dinámicamente los resultados de la evaluación según los tipos, de modo que ya no se aplique un estándar único. A través del análisis de correlación entre las puntuaciones de DWES para PC y las calificaciones de los expertos, los resultados son que, en comparación con Dr. Scratch, el coeficiente de correlación incrementó. Según Kazimoglu (2020), se necesitan nuevos métodos de motivación para llamar la atención de los estudiantes y adaptarse a sus nuevos patrones de aprendizaje ya que la forma en que las personas aprenden ha cambiado drásticamente en las últimas dos décadas. Para solucionar el problema, el autor propone como principal herramienta de aprendizaje los VJ. En su estudio, 151 de los 190 estudiantes completaron con éxito su VJ diseñado y los hallazgos sugieren que se hayan incrementado la: Motivación, percepción y confianza significativamente después de jugar el VJ.

Con base en Liu *et al.* (2023), las habilidades de PC generalmente se consideran una habilidad básica para la resolución de problemas que se van ampliando gradualmente filtrándose a grupos de edad más jóvenes. Diseñó un proyecto basado en un juego para investigar la motivación de aprendizaje de los estudiantes y el desempeño de VJ en diferentes entornos de programación de un total de 108 estudiantes de quinto grado. Los hallazgos sugieren que el aprendizaje basado en juegos puede mejorar sustancialmente el interés por aprender. Según Piaget, la Cuarta etapa de desarrollo de los estudiantes, llamada etapa de operaciones formales, está establecida a partir de los 12 años y durante toda su vida adulta (Eligeeducar, 2019). El estudio de Chang *et al.* (2023) es particularmente importante porque representa el primer intento sistemático de desarrollar ciencia, tecnología, ingeniería, artes y matemáticas (STEAM, por sus siglas en inglés), integrando el Aprendizaje Basado en Proyectos (ABP), como método de enseñanza transdisciplinario para fomentar la creatividad y las habilidades del PC de los estudiantes. Los resultados de este enfoque incluyen mejoras significativas en varias dimensiones de la vida de los estudiantes: Creatividad, fluidez, flexibilidad, originalidad, elaboración y habilidades PC mejoradas. El estudio involucró a 54 estudiantes de secundaria, alumnos de séptimo grado, con 28 estudiantes

en el grupo experimental y 26 estudiantes en el grupo de comparación. El grupo experimental demostró mejoras significativas en la creatividad mientras que el grupo de comparación no mostró cambios significativos.

El problema de la falta de competencias desarrolladas en PC no se reduce solo al área académica, a nivel industrial existen proyectos para realizar programación de robots industriales con lentes de AR. El proyecto de Guhl *et al.* (2017) es un desarrollo que permite programar los robots industriales de las líneas de las fábricas. Los cambios en estas líneas de producción son complejos y costosos además del tiempo que se tarda, esta innovación permite ejecutar los cambios sin estar presente y realizando pruebas de manera anticipada. Propuestas con AR orientados a la pedagogía especialmente a la programación como referencia se observan a niveles básicos como propuesta al desarrollo del PC como el caso de Gardeli y Vosinakis (2019), está diseñado para estudiantes de colegio de 9 a 10 años y un tablero que funciona como mapa y marcador al mismo tiempo, cada estudiante tiene las instrucciones y a su vez son las instrucciones que los estudiantes tienen para desarrollar la actividad, con el celular ellos capturan las instrucciones y observa las animaciones para obtener su resultado y verificar si fue correcto. El proyecto de Nikaido y Ventura (2016) es una estrategia de programación gráfica que facilita e induce a los nuevos estudiantes de sistemas.

Lo que se busca es incentivar a los alumnos a programar con retos pequeños incrementado la dificultad de manera progresiva, de esta manera los estudiantes encuentran el manejo de instrucciones sencillas, ciclos entre otros, en planos 2D en un programa para desarrollar la lógica de programación. Apaza *et al.*, (2019) realizan un trabajo donde propone un nuevo marco de diseño de contenidos de alto nivel de AR. El proyecto contó con la implementación

del framework usando el motor de creación de AR llamado UNITY y la extensión "Vuforia", para la creación de aplicaciones que usan "marcadores" (entendiendo marcadores, como los puntos de referencia sobre la imagen del entorno real para superponer una imagen digital AR). Los resultados de las pruebas de usabilidad basadas en ISO 9241-11 muestran que la propuesta es efectiva porque todos los usuarios pudieron completar todas las tareas, es eficiente porque los usuarios pudieran crear aplicaciones de AR en menos de 5 minutos y verificaron la satisfacción del usuario con una alta puntuación promedio en el cuestionario: *System Usability Scale*, cuyo resultado fue "altamente aceptable". Un estudio realizado por Yang (2019), propone un sistema de aprendizaje de robótica educativa virtual basado en AR, llamado: "AR Bot", que tiene como objetivo apoyar las actividades de aprendizaje de PC y educación: STEM. El sistema consta de dos elementos principales: La tarjeta AR Bot y la aplicación AR Bot. La tarjeta AR Bot es un activador de AR para mostrar materiales de aprendizaje y cuestionarios, se pueden usar teléfonos inteligentes con la aplicación AR Bot para escanear las tarjetas AR Bot y luego ingresar a las unidades de aprendizaje que utilizan el editor de programación visual en bloques para practicar diversas tareas del robot. Los movimientos del robot virtual se pueden demostrar en escenas auténticas con tecnología AR a través de dispositivos móviles. El sistema AR Bot está diseñado para facilitar PC y la educación STEM en escuelas K-12. Según Comber *et al.* (2019), aprender a codificar mediante la creación de VJ promete una gran motivación, compromiso y diversión para los estudiantes. En este estudio, se utilizó el entorno de desarrollo de juegos usando el motor de creación de AR: UNITY, una herramienta ampliamente difundida entre desarrolladores de juegos profesionales para enseñar a los estudiantes de secundaria a desarrollar sus propios VJ. La

investigación se basó en la enseñanza del diseño de VJ acompañado de una combinación de tutorías, explicaciones de los profesores, apoyo y colaboración entre los estudiantes. Los resultados confirmaron la hipótesis de que el desarrollo de VJ era atractivo para la mayoría de los jóvenes estudiantes.

Marco Teórico

El Pensamiento computacional (PC)

Rijo-García (2022) y Snalune (2015) coinciden en que el PC exige en el individuo el desarrollo de seis habilidades: 1) Descomposición: dividir el problema en partes más pequeñas. 2) Reconocimiento de patrones: encontrar similitudes dentro o fuera de la situación problemática. 3) Abstracción: simplificar un problema complejo centrándose solo en los detalles relevantes. 4) Crear un algoritmo: definir los pasos necesarios para resolver el problema. 5) Codificación: desarrollar un programa en un lenguaje que la computadora entienda según el algoritmo. 6) Depuración: corregir errores o mejorar el programa. El PC utiliza las funciones ejecutivas del ser humano para lograr su objetivo. Según Bauermeister (2008), las funciones ejecutivas del ser humano son actividades mentales complejas, necesarias para planificar, organizar, guiar, revisar, regularizar y evaluar el comportamiento necesario para adaptarse eficazmente al entorno y para alcanzar metas.

Contraste de las hipótesis para muestras dependientes

Según Tejedor y Murgiondo (2006), cuando a una muestra se le aplica un tratamiento y se desean contrastar las hipótesis para determinar si existe una diferencia significativa por causa de la aplicación del tratamiento a esa muestra, a este tipo de prueba se le conoce como: Prueba de hipótesis para muestras relacionadas o dependientes. La prueba "t" de Student para contrastar la diferencia significativa de las diferencias entre los datos de las muestras dependientes se calcula con la siguiente fórmula, según coinciden Dietrichson (2019) y Datalab (2023):

$$\text{Ec. 1. } t = \frac{\bar{x}_{diff} - 0}{s_{\bar{x}}} \quad \text{con: Ec. 2. } s_{\bar{x}} = \frac{\bar{s}_{diff}}{\sqrt{N}}$$

Donde las variables de entrada para las ecuaciones 1 y 2 se definen como: \bar{x} -diff = Promedio de las diferencias de los datos entre los grupos, antes y después del tratamiento N = Tamaño de la muestra, S_{diff} = Desviación standard de las diferencias de los datos antes y después del tratamiento, S_{x-} = Error típico estimado de la media. La prueba "t" de Student es válida para valores muestrales de hasta 30 datos en la mayoría de los casos, sin embargo, según algunos autores puede usarse con una muy buena aproximación a 60 datos, incluso las tablas contemplan valores de grados de libertad que equivalen al número de datos menos 1 hasta el infinito. Se adjunta la tabla 1 (parcial) de los valores de la distribución "t". Se puede usar en una hoja de cálculo de:

Google Docs usando la función: “INV.T.2C (0.01,37)” para aplicar un 95% de confianza, un 5% de significancia con $38 - 1 = 37$ grados de libertad para 38 datos recolectados

en las muestras pareadas, el valor crítico calculado en la hoja de cálculo es una interpolación entre los datos de $t = 50$ y $t = 30$ cuyo resultado es: 2,02.

Tabla 1. Valores (parciales) de “t” de Student para dos colas en función del valor de confianza

Valor de t para un intervalo de confianza de Valor crítico de t para valores de P de número de grados de libertad	90%	95%	98%	99%
	0.10	0.05	0.02	0.01
1	6.31	12.71	31.82	63.66
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

Cálculo de la muestra para una población finita

Según Vivanco (2005), la ecuación 3 puede ayudar a encontrar el tamaño de la muestra sobre una población finita para realizar un estudio sobre la misma:

$$\text{Ec. 3 } n = \frac{\frac{N}{N-1} pq}{\frac{e^2}{z^2_{\alpha/2}} + \frac{pq}{N-1}}$$

Donde: N es el tamaño de la población, p y q: son las probabilidades de éxito y fracaso de encontrar un individuo en la población con las características necesarias. Al ser desconocido este valor, puede asumirse $q = p = 0,5$ que maximiza el tamaño de la muestra con la probabilidad del 50% para cada caso e = Porcentaje de error esperado en los resultados obtenidos, se digita de forma decimal. Estos valores suelen ser generalmente menor o igual a $5\% = 0,05$, Z = Corresponde al valor de la variable normalizada de confianza en la campana de Gauss. El valor de Z está asociado al valor de confianza deseado, de forma común, suele usarse $Z = 1,96$ para una confianza del 95%.

Coefficiente de Kuder-Richardson

Cuando se desea medir una variable cualitativa que no es mensurable directamente tal como el nivel de inteligencia, capacidad de escritura, inclinación hacia la compra de un producto o nivel de PC, pueden usarse variables cuantitativas que si sean mensurables dentro de una prueba con el fin de determinar la variable cualitativa. Surge la necesidad de que las preguntas de la prueba diseñada deben estar lo más relacionadas entre sí para poder obtener al final una idea correcta de la medida de la variable latente y que los resultados obtenidos sean resultados fiables. Según Aiken (2003), existen 2 coeficientes importantes para evaluar el nivel de consistencia interna de la prueba diseñada según las respuestas obtenidas después de aplicar, estos coeficientes son el de Cronbach y Kuder-Richardson. Aunque el coeficiente de Cronbach es muy utilizado en estudios asociados a la educación como lo considera Cutumisu *et al.* (2022), Aiken (2003) propone que cuando una prueba realizada a estudiantes se caracteriza por tener preguntas cuyas respuestas son del tipo dicotómicas, es decir, binarias, falso o verdadero, correcto o incorrecto, se debe calcular el coeficiente

de Kuder-Richardson para determinar qué tan consistente es la prueba a nivel interno, es decir, qué tan relacionadas están las preguntas entre sí para que los resultados sean fiables y que la interpretación de los mismos sea acorde con el modelamiento de la realidad. El coeficiente de Cronbach es bastante útil para evaluar la fiabilidad interna de las pruebas que tienen varias opciones de respuestas usando la escala de Likert (Morales *et al.*, 2003), por otro lado, si las pruebas tienen opciones de respuestas dicotómicas (Correcto - incorrecto), se debe usar la fórmula de Kuder-Richardson para determinar la fiabilidad de la prueba. Cualquiera de los dos coeficientes calcula en porcentaje el nivel de fiabilidad. La fórmula de Kuder-Richardson también es conocida como “KR20” o “r20” y utiliza información de la dificultad de cada pregunta para poder ser calculado ya que se basa en el número de preguntas correctas contestadas en la prueba, además de tener en cuenta la suma de sus varianzas y la varianza total de las respuestas de la prueba.

A continuación, se escribe la fórmula del coeficiente de Kuder-Richardson para el cálculo de la fiabilidad interna de la prueba de respuestas dicotómicas:

$$\text{Ec. 4 } r_{20} = \left(\frac{K}{K-1} \right) \left(\frac{\sigma^2 - \sum pq}{\sigma^2} \right)$$

Donde: K = Es el número de preguntas del instrumento de evaluación (número de preguntas de la prueba), p = Porcentaje en fracción del número de personas que responden correctamente cada pregunta de la prueba. A cada pregunta de la prueba debe calcularse el porcentaje de personas que acertaron, q = Porcentaje en fracción del número de personas que responden correctamente cada pregunta de la prueba. Es importante que a cada pregunta de la prueba debe calcularse el porcentaje de personas que NO acertaron. Este valor puede calcularse como $1 - p$ para cada pregunta de la prueba. σ = Desviación estándar de los datos del instrumento de evaluación. En la mayoría de los estudios, aunque depende del investigador, suele atribuirse a un buen nivel de fiabilidad a valores del coeficiente de Kuder-Richardson superiores a 0,6, sin embargo, otros autores consideran excelente la fiabilidad para valores superiores a 0.7 hasta 1.0 (Cutumisu *et al.*, 2022). Puede asumirse una baja fiabilidad en la prueba para valores entre cero y 0.6 (Aiken, 2003). Puede utilizarse una hoja de cálculo para facilitar la determinación del valor del “r20”.

Metodología

Esta investigación tiene como objetivo determinar si el entrenamiento previo en PC puede lograr desarrollar competencias en los estudiantes que superan los doce años, dado que según el psicólogo y epistemólogo Jean Piaget, a partir de esta edad se está en la capacidad de realizar operaciones matemáticas formales de tal manera que mejoren sus resultados en pruebas certificadas. Operaciones

matemáticas formales como son: suma, resta, multiplicación, división, potencias, raíces y operaciones algebraicas avanzadas. Así mismo, Piaget establece que las habilidades matemáticas se definen como la capacidad de un estudiante para comprender, aplicar y utilizar conceptos, procedimientos y estrategias matemáticas para resolver problemas y razonar.

Capacitación previa en operaciones formales

Los 38 estudiantes con edades entre 17 y 20 años fueron seleccionados de manera aleatoria de una población de estudiantes del mismo nivel educativo (Primer semestre de ingenierías). Se asignaron a sus respectivos grupos de intervención y de control utilizando un proceso aleatorio. A la muestra de 38 estudiantes se les realizó, primero, una clase virtual de nivelación y repaso de las operaciones aritméticas básicas usando calculadora: Casio fx82Es o similar, esto con el fin de confirmar el dominio en su uso y el dominio de la solución de problemas que involucran operaciones básicas matemáticas como: adición, sustracción, multiplicación, división y cálculos de porcentajes usando reglas de 3. Esta muestra constituyó el grupo de control y se le aplicó una prueba de 5 preguntas que requieren PC.

Diseño del experimento

Con base en Domínguez y Castaño (2016), cualquier experimento que se vaya a realizar debe tener en cuenta 3 aspectos: factores, niveles y tratamientos. Los factores hacen referencia a la cantidad de variables independientes que intervienen en el experimento, los niveles hacen referencia al valor que se le van a asignar a cada variable dentro del experimento, los tratamientos hacen referencia a la mezcla de factores con los niveles. El experimento más sencillo sería el de 1 factor

con 2 niveles, conocido como experimento unifactorial. Los experimentos unifactoriales, en donde todos los errores experimentales que pueda haber por causa de otros factores que no se han tenido en cuenta por que se asume que sus aportes de error son muy pequeños, se les conocen como experimentos completamente aleatorizados (Dominguez, 2016). El autor propone que a estos experimentos se les puede modelar mediante estadística inferencial usando un estadístico de prueba para verificar con cierto nivel de confianza y significancia la hipótesis nula o la alternativa usando dos colas usando los valores de la distribución "t" de Student o bien, la distribución "z" si es el caso, que para un factor con dos niveles podrían definirse así: hipótesis nula: no hay efectos diferentes en los dos tratamientos. Hipótesis alternativa: los tratamientos tienen un valor significativamente diferente.

Diseño del cuestionario

Las preguntas tanto de la prueba 1 y las de la prueba 2, que son iguales en contenido, pero con diferente retórica, corresponden a 5 preguntas seleccionadas del examen de 50 preguntas que realiza el estado en Colombia en el área de matemáticas a los estudiantes de último año de educación media: SABER 11.

Este examen está validado y es homologable con otros exámenes de países tales como: Alemania, Australia, Bolivia, Brasil, Bulgaria, Canadá, Chile, Estados Unidos, España, México, Suiza, entre otros, según informa el MEN (2006). El reporte de gestión del MEN 2018-2022, entre las páginas 64 y 90, deja claro que el SABER 11 tiene por objetivo evaluar varias competencias del individuo, entre ellas, el PC (ICFES, 2018). Cuando se diseña un examen, deben considerarse múltiples variables que pueden influir en su duración y asignarle el número de preguntas correcto. Un estudio realizado por Trauzettel-Klosinski y Dietz

(2012) para 17 idiomas, incluido el español, determinó que la velocidad media de lectura comprensible para una persona de 10 años en adelante es de $1,42 \pm 0,13$ textos/min (\pm Desviación Estándar), 184 ± 29 palabras/min, 370 ± 80 sílabas/min y 863 ± 234 caracteres/min. De esta manera, puede estimarse el tiempo correcto por cada pregunta. Con base en estos datos, el MEN asigna para el SABER 11 aproximadamente 2,2 minutos por cada pregunta, sin embargo, en el diseño de nuestro cuestionario, se asignó tiempo indefinido con el fin de no presionar al estudiante y que se concentrará en seguir los algoritmos propuestos por cada pregunta.

Con relación al número adecuado de preguntas, los resultados de 2 estudios realizados por Sarrias-Ramis *et. al.* (2010) a estudiantes de 2 universidades de Barcelona, España, concluyen que la cantidad correcta de preguntas que debe tener una prueba para evaluar una asignatura completa debe ser de 100, sin embargo, el estudio mostró que la fiabilidad de los resultados no es significativamente diferente cuando los exámenes tienen la mitad de las preguntas. Su conclusión general es que la evaluación del rendimiento académico no dependería fundamentalmente del número de preguntas y que sería más importante buscar la representatividad y relevancia de éstas (Sarrias-Ramis *et. al.*, 2010).

En Colombia se diseñan las preguntas de los exámenes SABER 11 con base en MEN (2006). El MEN clasifica las competencias obtenidas por un estudiante desde grado 1 hasta grado 11 en 5 estándares básicos por competencias, la primera, grado 1 a 3, la segunda, grado 4 y 5, la tercera, grado 6 y 7, la cuarta, grado 8 y 9, quinta, grado 10 y 11. Un análisis realizado al documento del MEN (2006) muestra que, si se aplica la propuesta de 100 preguntas de Sarrias-Ramis *et. al.* (2010), un examen de estado en el ámbito de matemáticas debería contener 100 preguntas que evaluarán los 5 estándares de competencias, 20 preguntas por cada estándar, pero esto en la práctica sería muy extenso. Ya que la fiabilidad no disminuye significativamente con la mitad de las preguntas, el examen cuenta con 50 en total para matemáticas, 10 para cada competencia estándar. Esto puede apreciarse en los cambios que ha experimentado el examen de estado después del año 2000 en Colombia, pues ya no se llama examen ICFES sino SABER 11 y ya no se presenta los días sábado y domingo sino solo el domingo, naturalmente, posee aproximadamente la mitad de las preguntas que antes.

Con base en MEN (2006) puede verificarse que grados 8 a 11 desarrollan 2 de las 5 competencias estándar y dentro de estas 2, se desarrollan en total 56 competencias específicas que se agrupan en: Pensamiento numérico y sistemas numéricos nivel 1 y 2, Pensamiento espacial y sistemas geométricos nivel 1 y 2, Pensamiento métrico y sistemas de medidas nivel 1 y 2, Pensamiento aleatorio y sistemas de datos nivel 1 y 2 y Pensamiento variacional y sistemas algebraicos-analíticos nivel 1 y 2. De esta manera, el examen debe abordar 20 preguntas de temas relacionados desde el grado 8 al grado 11 y evaluar las 56 competencias específicas de

los temas impartidos, por esta razón, el SABER 11 mezcla competencias en una sola pregunta. Al analizar el listado de 56 competencias específicas del examen puede verse que 14 están relacionados con el PC, aplicando una regla de 3, se explica por qué la prueba de SABER 11 2018 utilizó 5 preguntas para evaluar el PC de los estudiantes en simultáneo a otras competencias.

Condicionamiento previo del estudiante a la prueba 1

Al utilizar como modelo de cálculo a la estadística inferencial asumiendo la evaluación de la hipótesis a dos colas, esta investigación asume que el grupo de control no contaba con ningún entrenamiento previo en PC, sin embargo, debe tenerse en cuenta que así como dentro del grupo pueden haber estudiantes que presentaron su examen de estado hace muchos años (en tiempos donde el PC no se evaluaba), es de esperarse que la mayoría de los estudiantes del grupo haya presentado recientemente el examen de estado SABER 11, probablemente en 2023, esto significa que NO se puede asegurar que los estudiantes tengan o no un condicionamiento previo implícito. Como el último examen SABER 11 del año en Colombia se presenta en el mes de agosto, en este caso, agosto de 2023, y esta investigación se realizó en enero de 2024, se limita la validez del preconditionamiento nulo en preguntas que incluyen PC para el grupo de control en los últimos 4 meses.

Condiciones de ejecución del experimento

Las prepruebas y postpruebas se realizaron en condiciones controladas. Los estudiantes

estuvieron bajo supervisión durante todo el proceso para garantizar que no recibieran ayuda externa. La prueba 1 se realizó con las siguientes condiciones: el estudiante debe resolver todas las preguntas sin límite de tiempo. El entorno de la prueba debió permitir que el estudiante respondiera de manera concentrada (sin distractores como celulares, mascotas, peluches, almohadas, ruidos externos, otras personas, elementos de oficina que sean inútiles, entre otros), que no seleccionara respuestas al azar y que respondiera según considerase correcto a partir de sus análisis y conocimientos. Podía usar calculadora tipo Casio fx-82ES o similar y una hoja en blanco. Se verificaron calculadoras con suficiente batería y lapicero con suficiente tinta. Se utilizó un cronómetro para cada estudiante y se daba inicio en el momento justo que comenzaba a leer la primera pregunta, la finalización del conteo se realizó justo en el momento que el estudiante terminaba la última pregunta de la prueba.

Diseño del Video Juego (VJ) de Realidad Aumentada (AR) para el desarrollo del Pensamiento Computacional (PC) del estudiante después de aplicarle la prueba 1

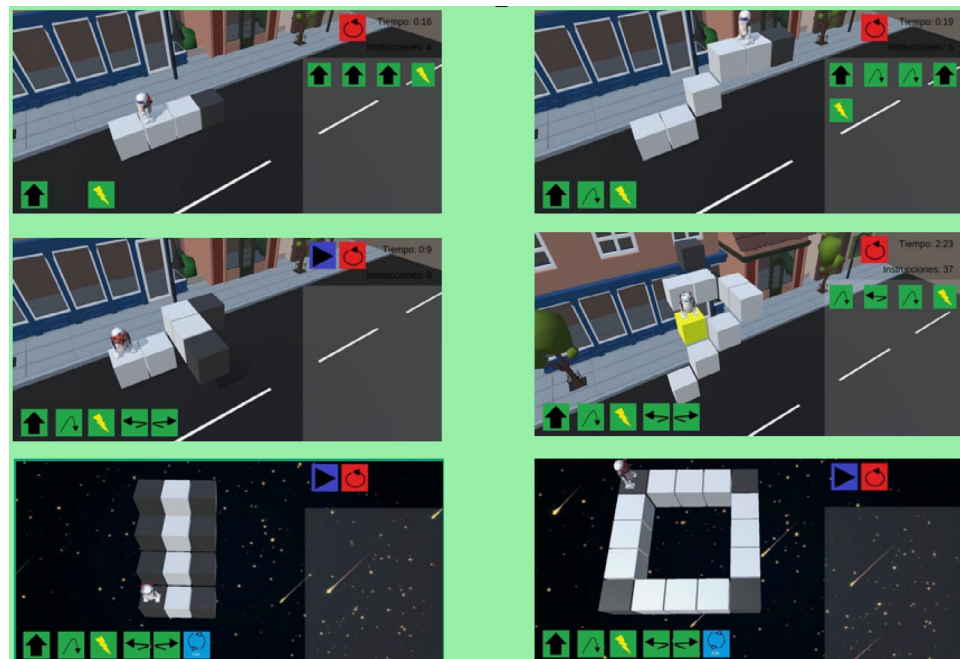
Este trabajo de investigación incluye la creación de un juego de AR programado para Android y Windows, generado con el motor de videojuegos: UNITY, el cual puede ejecutarse en un OCULUS QUEST 2, descargable en celular. El VJ consiste en entrenar en pensamiento computacional (PC) al estudiante, realizando una serie de pasos hasta lograr una tarea principal que se le asigna en el juego.

Figura 1. Captura de imagen de introducción Videojuego de PC diseñado para el tratamiento de PC



El juego enseña principios básicos de programación de forma divertida e interactiva. Los jugadores controlan un personaje en un entorno 3D isométrico, usando secuencias, bucles y funciones para resolver desafíos. Con niveles progresivos y disponible en varias plataformas, es una herramienta popular para aprender lógica de programación y pensamiento algorítmico. Después de la prueba 1 se aplicó 10 veces consecutivas el juego al grupo de control. Se esperaron 30 minutos para abordar la prueba 2.

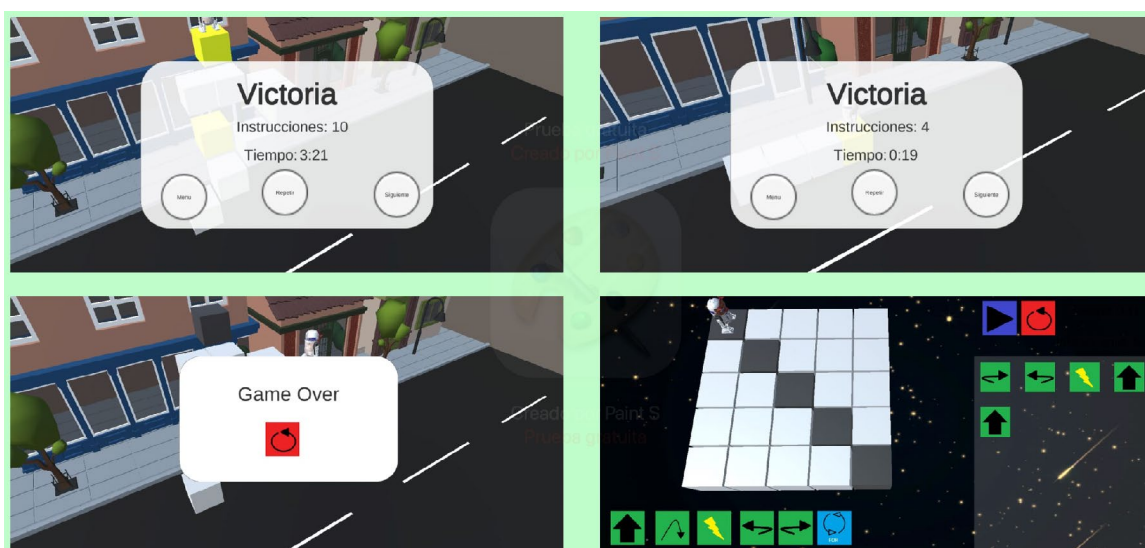
Figura 2. Captura de imágenes de los diferentes escenarios del Videojuego de PC, en total son 25 escenarios



El proyecto tendrá 3 escenarios, cada uno con tres niveles de complejidad, los marcadores indicarán el nivel y el escenario que se desea jugar para un total de 9 marcadores. Estos niveles de complejidad están relacionados con las etapas de razonamiento para facilitar la introducción a la programación. Las preguntas de la segunda prueba fueron

cambiadas de orden y se escribieron palabras diferentes con el fin que el estudiante no descubriera que se trataba de la misma prueba 1. La pregunta 3 no tuvo ninguna variación con respecto a la de la prueba 1. Con la prueba 2 se pudieron recolectar nuevamente datos de tiempo y efectividad en la respuesta correcta a las preguntas y se analizaron los resultados.

Figura 3. Escenarios y módulos del Videojuego. Los aciertos permiten continuar con el siguiente nivel, de lo contrario el jugador se enfrenta al mismo nivel con cambios en el escenario



Limitaciones de la investigación

Con el ánimo de ser lo suficientemente rigurosos, se plantean las siguientes limitaciones:

Los estudiantes contaban con diferentes niveles de exposición previa al pensamiento computacional y al uso de videojuegos educativos, lo que podría haber afectado su desempeño en las pruebas.

Aunque los resultados de esta investigación son estadísticamente significativos, la muestra de 38 estudiantes es relativamente pequeña. Se pueden necesitar más estudios con muestras más grandes para confirmar estos hallazgos en diferentes contextos educativos.

Resultados y análisis

Se pudieron verificar diferencias negativas de tiempo entre las dos pruebas para cada pregunta. Los datos muestran que el promedio de tiempo de respuesta por pregunta en los 38 estudiantes evaluados fue de 1957,26 segundos antes de la capacitación en PC, y posterior a ella disminuyó a 1165,28, presentándose una diferencia de -791,97 segundos para la segunda prueba. A partir de los datos recolectados, se realizó la sumatoria de los tiempos promedio de cada pregunta para obtener el valor total del tiempo empleado por los 38 estudiantes para la prueba 1 y de la misma manera para el tiempo de la prueba 2. Estos

tiempos se muestran en el diagrama de barras de la Figura 4. En la Figura 5, se presenta un diagrama de barras más específico con los tiempos de respuesta promedio para cada pregunta de las pruebas 1 y 2 antes y después del tratamiento del Pensamiento Computacional (PC) con el Video Juego (VJ), respectivamente, en donde puede verificarse que los estudiantes lograron en general disminuir el tiempo empleado para resolver cada pregunta.

Los datos de la Figura 4 sugieren una diferencia de medias de tiempo empleado de $X_{diff} = -263,99$ segundos con una desviación estándar de: $S_{diff} = 271,23$ segundos para todos los datos de diferencia de medias de tiempo empleados para responder cada pregunta antes y después del tratamiento. Los datos fueron procesados en una hoja de cálculo utilizando Google Docs y el resumen de los análisis se presenta en la Tabla 1. Es interesante destacar que el número promedio de respuestas correctas antes de aplicar el tratamiento de PC era del 60%, luego, después de jugar el VJ, se pudo verificar que el promedio de los 38 estudiantes se incrementó en un 15% en el total de respuestas correctas. Estos resultados se exponen en la Figura 4.

Figura 4. Sumatoria de los tiempos totales promedio empleados para dar respuesta a la prueba 1 y 2 antes y después del tratamiento de PC con el videojuego

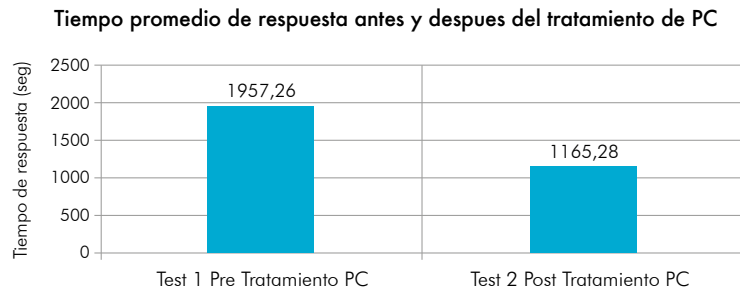


Figura 5. Tiempos de respuesta promedio a cada pregunta antes y después de aplicar el tratamiento de desarrollo de competencias de PC usando el videojuego de la Figura 1

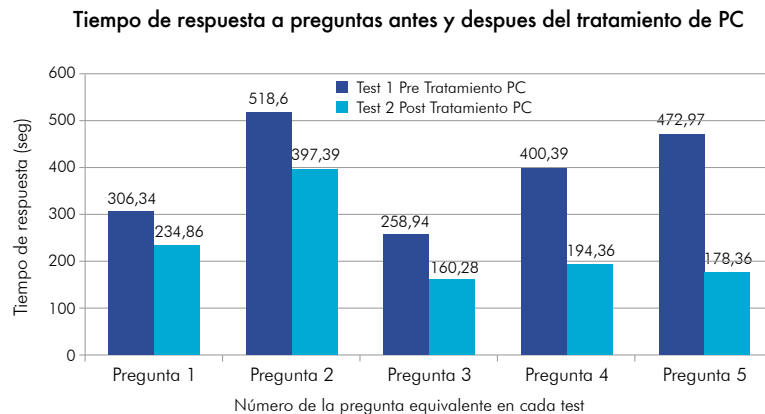
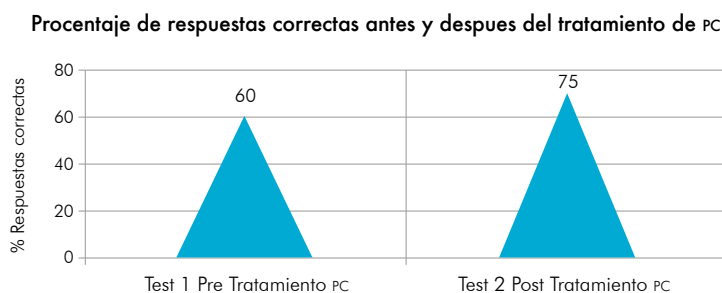


Figura 6. Promedio de respuestas correctas para 38 estudiantes de las pruebas 1 y 2 antes y después del tratamiento de PC con el videojuego



La correlación entre el aumento en las respuestas correctas y la reducción del tiempo de respuesta evidencia una mejora en la capacidad de razonamiento, desarrollo del PC y resolución de problemas, lo que sugiere una mejora en la capacidad para resolver problemas algorítmicos, de una manera más rápida y efectiva.

Utilizando las ecuaciones 2 y 1 respectivamente para determinar: S_x : el error típico estimado de la media y luego el estadístico de prueba "t", puede verse que el primero tiene como resultado: 121,3 segundos y el estadístico un valor de: 2,17 adimensional. Este valor de 2,17 fue comparado con el dato del valor crítico de la distribución de "t" Student, el cual puede encontrarse por interpolación de los datos de la tabla 1 o bien, usando la función: " $=INV.T.2C(0.05,37)$ " en la hoja de cálculo para determinar el valor exacto del valor crítico "t". Los cálculos del valor crítico se obtuvieron asumiendo un 95% de confianza, en consecuencia, 5% de significancia para un total de 38 estudiantes encuestados lo que sugiere $n - 1 = 38 - 1 = 37$ grados de libertad.

Con estos parámetros, se obtuvo 2,02 como valor crítico que es un número menor a 2,17 del valor "t" de Student, por lo cual, sugiere que la hipótesis nula SEA RECHAZADA ya que el resultado de la "t" se encuentra en la zona de

rechazo de la distribución de Student. Las hipótesis planteadas fueron: hipótesis nula: No hay efectos diferentes en los dos tratamientos; hipótesis alternativa: Los tratamientos tienen un valor significativamente diferente. Esto sugiere aceptar la hipótesis alternativa con una confianza del 95% en que sí existe una diferencia significativa entre las medias de los tiempos empleados para resolver las preguntas de la prueba después de aplicar un tratamiento de desarrollo del PC usando el vJ de la Figura 1. De la misma manera, puede decirse que con una confianza del 95% existe un incremento significativo del 15% en el número de respuestas correctas que el estudiante puede lograr aplicando el tratamiento de PC con el vJ. Para obtener el resultado de la ecuación 4, se utilizó una hoja de cálculo de Google Docs, esto con el fin de determinar el índice de Kuder-Richardson. Según Cutumisu *et al.* (2022), en la mayoría de los estudios suele atribuirse a un buen nivel de fiabilidad a valores del coeficiente de Kuder-Richardson superiores a 0,6 y según Aiken (2003) se considera una excelente fiabilidad los valores superiores 0.7 hasta 1.0. Para esta investigación, el valor obtenido para la fiabilidad y coherencia de la prueba 1 fue de 0,62 el cual puede asumirse como aceptable, esto abre la posibilidad a estudios futuros que usen preguntas en pruebas más compactas para aumentar la fiabilidad de los resultados.

Tabla 1. Resumen de datos analizados

	Tiempo Respuesta Prueba 1 Pre Entrenamiento ct (seg)	Tiempo Respuesta Prueba 2 Post Entrenamiento ct (seg)	Diferencias de tiempo X1 - X2 (seg)	
Pregunta 1	306,34	234,86	71,47	
Pregunta 2	518,60	397,39	121,21	
Pregunta 3	258,94	160,28	98,65	
Pregunta 4	400,39	194,36	206,02	
Pregunta 5	472,97	178,36	294,60	

Variable	Antes del tratamiento	Después del tratamiento	Diferencia	Desviación estándar
Total tiempo promedio empleado en la prueba (seg)	1957,26	1165,28	791,97	271,23
Respuestas Correctas (%)	60%	75%	15%	
Sx= Error típico estimado de la media				121,3
Estadístico de prueba "t" de Student				2,17
Valor crítico para t				2,02

Fuente: elaboración propia.

Conclusiones

Se diseñó un Videojuego (VJ) educativo con mecánicas de interacciones basadas en Realidad Aumentada (AR) para el entrenamiento de habilidades de razonamiento y desarrollo del Pensamiento computacional (PC) del estudiante. El VJ fue diseñado con el motor UNITY y se integró el aprendizaje basado en proyectos (ABP) como método de enseñanza para fomentar la creatividad y las habilidades del PC de los estudiantes.

Tras aplicar una primera prueba con preguntas que requerían el uso de PC a 38 estudiantes, los resultados mostraron, con un nivel de confianza del 95 %, un incremento del 15 % en su capacidad para resolver problemas con el uso del PC. Además, se observó una disminución significativa en el tiempo promedio de la prueba, que pasó a ser de 13.2 minutos menos en promedio en la segunda prueba, realizado con preguntas del mismo tipo. La validación de los efectos de la implementación del VJ se llevó a cabo mediante el coeficiente de Kuder-Richardson, obteniendo un valor de 0.62, lo que evidencia una consistencia interna aceptable de la prueba.

Con los resultados obtenidos, se propone a los educadores experimentar con la integración de video juegos educativos en los planes de estudio, con el objetivo de mejorar el pensamiento computacional y analizar el impacto del uso continuo de estas herramientas a lo largo de varios semestres académicos,

incluso en otras asignaturas que al igual requieren pensamiento lógico y algorítmico como la física. De igual forma analizar, de qué forma, al desarrollar el pensamiento crítico y la resolución de problemas, puede mejorar la creatividad y el trabajo en equipo en el ámbito profesional de los estudiantes.

Si bien una muestra de 38 estudiantes es interesante para la población estudiantil de la universidad, no es una muestra representativa dentro de las metodologías de investigación. Los resultados muestran una mejora significativa en la reducción del tiempo de respuesta en 13.2 minutos promedio y un incremento de las respuestas correctas de un 15%. Como recomendación para futuras investigaciones se propone una base de estudiantes superior que involucre una mayor cantidad de semestres con características similares, para mostrar los efectos positivos de los video juegos en el pensamiento lógico y algorítmico.

Agradecimientos

Este trabajo se realizó en el marco de la Maestría en Tecnologías de la Información Aplicadas a la Educación y del proyecto de investigación DTE-661-24 de la Universidad Pedagógica Nacional.

Referencias

- Aiken, R. L. (2003). *Tests Psicológicos y Evaluación*. Pearson.
- Apaza, Y., Paz, A. y Corrales, C. (2019). *SimpleAR: Augmented Reality high-level content design framework using visual programming* [Ponencia]. 38th International Conference of the Chilean Computer Science Society (SCCC), Concepción, Chile. <https://pure.unsa.edu.pe/es/publications/simplear-augmented-reality-high-level-content-designframework-us-2>

- Araujo, C., Lima, L. V. O. y Henriques, P. R. (2019). *An ontology-based approach to teach computational thinking* [Ponencia]. International Symposium on Computers in Education (SIIE), Tomar, Portugal. <http://doi.org/10.1109/siie48397.2019.8970131>
- Chai, X., Sun, Y., Luo, H. y Guizani, M. (2019). DWES: A dynamic weighted evaluation system for Scratch based on computational thinking. *Emerging Topics in Computing*, 10(2), 917-932. <https://ieeexplore.ieee.org/document/9321708>
- Comber, O., Motschnig, R., Mayer, H. y Haselberger, D. (2019). *Engaging students in computer science education through game development with Unity* [Ponencia]. 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates. <https://ieeexplore.ieee.org/document/8725135>
- Csernoch, M., Biró, P. y Máth, J. (2021). Developing computational thinking skills with algorithm-driven spreadsheeting. *IEEE Access*, 9(1), 153943-153959. <https://www.researchgate.net/publication/356011457>
- Datalab. (2023). *Prueba t para muestras relacionadas*. Datalab.es. <https://datatab.es/tutorial/paired-t-test>
- Domínguez, J. y Castaño, E. (2016). *Diseño de experimentos*. Alfaomega.
- Eligeeducar. (2019). *Según Piaget, estas son las 4 etapas del desarrollo cognitivo*. Eligeeducar.cl. <https://eligeeducar.cl/acerca-del-aprendizaje/segun-jean-piaget-estas-son-las-4-etapas-deldesarrollo-cognitivo/>
- Guhl, J., Tung, S. y Kruger, J. (2017). *Concept and architecture for programming industrial robots using augmented reality with mobile devices like Microsoft HoloLens* [Ponencia]. 22nd IEEE International Conference on

Emerging Technologies and Factory Automation (ETFA), Limassol. <https://ieeexplore.ieee.org/document/8247749>

ICFES. (2018). *Transformación por la calidad educativa del país*. Instituto Colombiano para la Evaluación de la Calidad de la Educación. <https://www2.icfes.gov.co>

Kazimoglu, C. (2020). Enhancing confidence in using computational thinking skills via playing a serious game: A case study to increase motivation in learning computer programming. *IEEE Access*, 8(1), 221831-221851. <https://www.x-mol.net/paper/article/1340049860763238400>

Liu, H., Wu, Z., Lu, Y. y Zhu, L. (2023). Exploring the balance between computational thinking and learning motivation in elementary programming education: An empirical study with game-based learning. *IEEE Transactions on Games*, 15(1), 95-107. <https://ieeexplore.ieee.org/document/9684681>

Mineducación. (2021). *Estadísticas de deserción y permanencia en educación superior SPADIES 3.0, indicadores 2021*. Ministerio de Educación Nacional. <https://www.mineduacion.gov.co/sistemasinfo/spadies/secciones/Estadisticas-de-desercion/>

Ministerio de Educación Nacional (MEN). (2006). *Estándares básicos de competencias en lenguaje, matemáticas, ciencias y ciudadanas*. https://www.mineduacion.gov.co/1621/articles-116042_archivo_pdf.pdf

Vivanco, M. (2005). *Muestreo Estadístico: Diseño y aplicaciones*. Editorial Universitaria.