

Can artificial intelligence (AI) help in computer science education? A meta-analysis approach

¿Puede ayudar la inteligencia artificial (IA) en la educación en ciencias de la computación? Un enfoque metaanalítico

Ahmed TLILI. Associate Professor. Smart Learning Institute of Beijing Normal University, China (ahmedtlili@ieee.org).

Abstract:

Several studies have investigated the effect of Artificial Intelligence (AI) on students' learning achievement in education. However, limited research targeted Computer Science (CS) education, which is considered crucial regardless of the future profession. Consequently, scant information exists on how AI might impact students' learning achievement in CS education. To address this research gap, this study conducts a systematic review and a meta-analysis to investigate how AI integration affects learning achievement in CS education and the potential moderating variables of this effect. Specifically, 28 studies ($n = 2765$ participants in total) were included and meta-analyzed, and the obtained effect size was very large ($g = 1.36, p < .001$). Particularly, intelligent tutoring systems (ITSs) were found to have the highest effect ($g =$

1.45; huge) as an AI technology. Additionally, the AI intervention duration and the geographical distribution of students are found to moderate the AI effect in CS education. The findings of this study can serve as a reference for various stakeholders (e. g., educators, computer scientists, instructional designers) on how to integrate AI and improve learning experiences and outcomes in CS education.

Keywords: computer science, computing, artificial intelligence, education, learning, collaborative intelligence, meta-analysis, learning achievement.

Resumen:

Varios estudios han investigado el efecto de la inteligencia artificial (IA) en los logros de aprendizaje de los alumnos

Date of receipt of the original: 2024-06-30.

Date of approval: 2024-08-06.

Please, cite this article as follows: Tlili, A. (2024). Can artificial intelligence (AI) help in computer science education? A meta-analysis approach [¿Puede ayudar la inteligencia artificial (IA) en la educación en ciencias de la computación? Un enfoque metaanalítico]. *Revista Española de Pedagogía*, 82 (289), 469-490. <https://doi.org/10.22550/2174-0909.4172>

en educación. Sin embargo, hay un volumen limitado de investigación enfocada en la educación en ciencias informáticas o de la computación (CC), un área que se considera crucial con independencia de la profesión futura. En consecuencia, existe poca información sobre cómo la IA podría influir en los logros de aprendizaje de los alumnos en la educación en CC. A fin de llenar este vacío en la literatura, este estudio realiza una revisión sistemática y un metaanálisis para investigar cómo la integración de la IA afecta a los logros de aprendizaje en la educación en CC y las posibles variables moderadoras de este efecto. En concreto, se incluyeron y metaanalizaron 28 estudios ($n = 2765$ participantes en total) y el tamaño del efecto obtenido fue muy grande ($g = 1.36$, $p < 0.001$). En particular, se ha hallado

que los sistemas de tutoría inteligente (STI) presentan el mayor efecto ($g = 1.45$, un efecto enorme) como tecnología de IA. Adicionalmente, se ha encontrado que la duración de la intervención de la IA y la distribución geográfica de los alumnos moderan el efecto de la IA en la educación en CC. Los hallazgos de este estudio pueden servir como referencia para las diversas partes interesadas (por ejemplo, educadores, científicos computacionales y diseñadores formativos) sobre cómo integrar la IA y mejorar las experiencias y los resultados de aprendizaje en la educación en CC.

Palabras clave: ciencias informáticas, computación, inteligencia artificial, educación, aprendizaje, inteligencia colaborativa, metaanálisis, logros de aprendizaje.

1. Introduction

1.1. Artificial intelligence in computer science education

The increased need to educate students about computer science (CS) has been highlighted by former U.S President Obama, who initiated “CSforAll” to provide all K–12 children CS education in the United States (Smith, 2016). After that, several initiatives have been launched to teach CS or a particular technology, such as Artificial Intelligence (AI), to all students. For instance, the Computer Science Teachers Association (CSTA) and the Association for the Advancement of Artificial

Intelligence (AAAI) announced in 2018 the establishment of national guidelines about teaching K-12 students AI. These guidelines revolve around “five big ideas” in AI, namely Perception, Representation & Reasoning, Learning, Natural Interaction, and Societal Impact (AAAI, 2018; Touretzky et al., 2023). Teaching students CS and the use of technology is becoming a crucial competence for everyday life and any future profession.

However, learning CS subjects might be difficult and require more cognitive tasks than other subjects (Tlili et al., 2015). Silva et al. (2019) also pointed

out the challenges of teaching CS when analyzing the different disciplines in the area. Particularly, the difficulties that students face when learning about algorithms and data structure are of major concern for educators worldwide (Silva et al., 2019). In the same vein, Waraich (2004) further highlighted that students find some computer science subjects like binary arithmetic and logic gates to be “dry” and “not very interesting”.

AI, on the other hand, has also been reported to support education by providing personalized learning content, recommendation, and predictions (Onesi-Ozigagun et al., 2024). This has triggered researchers to integrate AI as a technology to facilitate CS education. For instance, Georgia Institute of Technology integrated AI-based instructions to teach its online master students about AI, which is part of a computer science program (Goel & Joyner, 2017). Zhang et al. (2021) developed an AI-based interactive e-book to teach K-12 students about AI. Their findings revealed that students have a high acceptance level of this e-book for learning. Gerdes et al. (2017) developed an adaptable programming tutor for learning the Haskell programming language.

1.2. Research gap and study objectives

Despite the increase adoption of AI in education generally and in CS particularly, the effects of AI on students' learning achievement are still debatable (Kim & Lee, 2023). Several meta-analy-

sis studies were conducted in the literature to investigate the effects of AI on learning achievement (Lin et al., 2022; Zheng et al., 2023). However, these meta-analyses targeted education in general. Particularly, scant information exists on the effect of AI on students' learning achievement in CS education. In this context, several studies conducted a systematic review of the use of Intelligent Tutoring Systems (ITS) in CS education (Crow et al., 2018; Francisco & de Oliveira Silva, 2022). However, these studies only conducted a qualitative analysis and did not discuss learning achievement. Therefore, the impact of AI on learning achievement remains unknown.

One study by Nesbit et al. (2014) conducted a meta-analysis of using ITSs in CS education, where 22 articles published between 1998 and 2013 were analyzed. The overall effect size was moderate ($g = .46$). While this study (Nesbit et al., 2014) is considered one of the early attempts to discuss the present research topic, its results cannot be generalized because: of two reasons. First, it focused only on one type of AI technology (i. e., ITS) rather than covering all AI technologies. Therefore, the obtained results do not reflect the AI effect in general. Second, the articles covered were before 2013 (with one study in 2013), hence the findings can be considered outdated due to the rapid evolution of technology generally and AI particularly, where new waves are seen like generative AI (Tlili et al., 2023), among others.

To the best of our knowledge, no research conducted an analysis to investigate the effects of AI on learning achievement in CS education. Therefore, the question of whether AI as computer science technology can help in CS education remains open. While there is no agreed-upon definition of computer science, this study considers CS as the study of computers and of the phenomena connected with computing, notably algorithms, programs, and programming (Shaw, 1985). To address the highlighted research gap, this study answers the following research questions:

RQ1. What is the effect of AI on students' learning achievement in computer science education?

RQ2. Which variables moderate the effect of AI on students' learning achievement in computer science education?

2. Methodology

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines were followed (Page et al., 2021) to conduct this meta-analysis.

2.1. Data selection

A search was conducted in the following electronic databases: IEEE Xplore, Science Direct, Scopus, Taylor & Francis, and Web of Science. These databases were chosen because they are popular in the field of educational technology. The search timeframe was set starting

from 2011 because: (1) this year was considered the year where AI applications became more mature (Wang et al., 2023); and (2) since then, the research on the use of AI technologies in CS education has started increasing significantly (Nesbit et al., 2015). The search keywords were adapted from several AI in education reviews in the literature (e.g., Nesbit et al., 2015; Zawacki-Richter et al., 2019; Zheng et al., 2023) and are as follows: ((Artificial intelligence substring) AND (Computer science substring)), where:

- Artificial intelligence substring: artificial intelligence OR AI OR machine intelligence OR machine learning OR natural language processing OR deep learning OR robotic.
- Computer science substring: computer science OR technology OR computer literacy OR information and communication technology OR ICT.

All publication types (conference proceedings, journal articles, etc.) were considered because proceeding papers are the most common format for researchers in computer science and software engineering, unlike in the education field, where journal publication is more common (Nesbit et al., 2015).

The final search was conducted on January 1, 2024, and the whole process yielded 548 potential studies. After

removing duplicates, 205 potential studies were identified and put through the inclusion/exclusion criteria. A study was included if it (1) was in English, (2) was empirical research, (3) used AI for CS education, (4) was not qualitative or review research, (5) provided sufficient information (e. g., mean, median, and standard deviation) to calculate the effect size, and (6) included a control condition. Finally, 28 studies (2765 participants in total) were considered for this meta-analysis.

2.2. Meta-analysis

Comprehensive Meta-Analysis V.4 (Borenstein, 2022) software was used to conduct the current meta-analysis. In addition, Hedges' g was used to calculate the effect sizes (Hedges, 1981). The motivation behind using Hedges' g instead of Cohen's d effect size was that the differential sample size between studies may bias the estimated effect size. This bias affects studies with a sample size smaller than 20, in which case Hedges's g presents more reliable estimates than Cohen's d (Hedges & Olkin, 1985). Table 1 presents the 28 included studies in this present meta-analysis.

Three methods were used to assess publication bias. Firstly, the trim-and-fill method, with the intention of identifying publication bias by means of a funnel plot wherein the studies are represented by dots. If the dots are distributed on both sides of a vertical line, representing the average effect size, it is assumed that there is no publication

bias (Borenstein et al., 2010). Secondly, Rosenthal's (1979) fail-safe number aims to determine the number of studies with nonsignificant results of unpublished data needed to nullify the mean effect size. A fail-safe number larger than $5k+10$ (where k is the original number of studies included in the meta-analysis) is robust. This means that the effect size of unpublished studies is not likely to affect the average effect size of the meta-analysis. However, this method assumes that the mean effect size in the missing studies is zero (Borenstein et al., 2021). The third method was Egger's regression test, where a significant intercept suggests publication bias (Lin et al., 2018).

2.3. Coding scheme

To minimize the potential for bias, an online electronic data extraction form was designed (Kitchenham & Charters, 2007). This form contains the following information, which could help in answering the aforementioned research questions: (1) AI technology: it refers to the type of AI technology used during the learning process; (2) field of education: it refers to the educational domain where AI was used; (3) level of education: it refers to the academic level at which AI was used; (4) learning mode: it refers to where and how the learning process occurred; (5) intervention duration: it refers to the time length over which AI was used in education; and (6) geographical distribution: it refers to the geographical distribution of students who were involved in the learning process.

TABLE 1. Descriptive and inferential statistics of the 28 included studies.

| # | Author(s) & year | AI type | Level of education | Learning mode | Intervention duration | Geographical distribution | Hedges's g | Standard error | Sample | | |
|----|--------------------------------|--------------|--------------------|---------------|-----------------------|---------------------------|--------------|----------------|------------------------|-------------------|-------|
| | | | | | | | | | Number of experimental | Number of control | Total |
| 1 | Hayashi (2019) | ITS | Primary | Online | One month | Asia | 1.335 | 0.362 | 18 | 18 | 36 |
| 2 | Cuong et al. (2018) | ITS | Higher education | Online | One semester | Asia | 1.297 | 0.255 | 37 | 36 | 73 |
| 3 | Yin et al. (2021) | Chatbot | Higher education | Online | One semester | Asia | 0.000 | 0.200 | 51 | 48 | 99 |
| 4 | Perikoset (2017) | ITS | Higher education | Online | One month | Europe | 3.381 | 0.207 | 113 | 113 | 226 |
| 5 | Tegos and Demetriadis (2017) | Chatbot | Higher education | Online | One semester | Europe | 0.859 | 0.244 | 38 | 34 | 72 |
| 6 | Tegos et al. (2015) | Chatbot | Higher education | Online | One week | Europe | 0.903 | 0.315 | 21 | 22 | 43 |
| 7 | Tegos et al. (2016) | Chatbot | Higher education | Online | One week | Europe | 0.594 | 0.252 | 32 | 32 | 64 |
| 8 | Fidan and Gencel (2022) | Chatbot | Higher education | Online | One month | Europe | 9.320 | 0.711 | 54 | 40 | 94 |
| 9 | Abbasi et al. (2019) | Chatbot | Higher education | Online | One semester | Asia | 1.214 | 0.206 | 55 | 55 | 110 |
| 10 | Abbasi and Kazi (2014) | Chatbot | Higher education | Online | One semester | Asia | 1.852 | 0.280 | 36 | 36 | 72 |
| 11 | Grivokostopoulou et al. (2017) | ITS | Higher education | Online | One month | Europe | 1.929 | 0.140 | 150 | 150 | 300 |
| 12 | Essel et al. (2022) | Chatbot | Higher education | Online | One semester | Africa | 5.592 | 0.536 | 34 | 34 | 68 |
| 13 | Lin and Chen (2020) | Personalized | Higher education | Blended | One month | Asia | 1.337 | 0.223 | 48 | 49 | 97 |

| # | Author(s) & year | AI type | Level of education | Learning mode | Intervention duration | Geographical distribution | Hedges's g | Standard error | Sample | | |
|----|-------------------------------|--------------|--------------------|---------------|-----------------------|---------------------------|--------------|----------------|------------------------|-------------------|-------|
| | | | | | | | | | Number of experimental | Number of control | Total |
| 14 | Stamper et al. (2013) | ITS | Higher education | Online | One semester | North America | 0.086 | 0.140 | 105 | 98 | 203 |
| 15 | Ma et al. (2020) | ITS | Secondary | Online | One month | Asia | 4.777 | 0.455 | 38 | 36 | 74 |
| 16 | Su (2020) | ITS | Higher education | Online | One month | Asia | 0.598 | 0.226 | 40 | 40 | 80 |
| 17 | Chang et al. (2016) | ITS | Higher education | Blended | One semester | Asia | 0.552 | 0.188 | 58 | 58 | 116 |
| 18 | Lai et al. (2021) | Personalized | Higher education | Online | One semester | Asia | 1.726 | 0.254 | 43 | 41 | 84 |
| 19 | Eryilmaz and Adabashi (2020) | ITS | Higher education | Online | One semester | Europe | 0.342 | 0.183 | 60 | 60 | 120 |
| 20 | Arsovic and Stefanovic (2020) | Personalized | Higher education | Online | One semester | Europe | 0.628 | 0.161 | 81 | 79 | 160 |
| 21 | Winkler et al. (2020) | Chatbot | Higher education | Online | One semester | Europe | 0.682 | 0.240 | 37 | 35 | 72 |
| 22 | Song and Kim (2021) | Chatbot | Higher education | Online | One semester | North America | 0.553 | 0.269 | 27 | 29 | 56 |
| 23 | Kumar (2021) | Chatbot | Higher education | Online | One semester | Asia | 0.243 | 0.256 | 30 | 30 | 60 |
| 24 | Farah et al. (2022) | Chatbot | Higher education | Online | One month | Europe | 0.205 | 0.432 | 11 | 9 | 20 |
| 25 | Al-Abdullatif et al. (2023) | Chatbot | Higher education | Blended | One semester | Asia | 0.338 | 0.257 | 30 | 30 | 60 |
| 26 | Ortega-Ochoa et al. (2024) | Chatbot | Higher education | Online | One semester | Europe | 0.243 | 0.143 | 101 | 95 | 196 |
| 27 | Sun et al. (2024) | Chatbot | Higher education | Online | One semester | Asia | 0.306 | 0.220 | 43 | 39 | 82 |
| 28 | Ahn and Oh (2024) | ITS | Higher education | Blended | One semester | Asia | 0.459 | 0.314 | 20 | 20 | 40 |

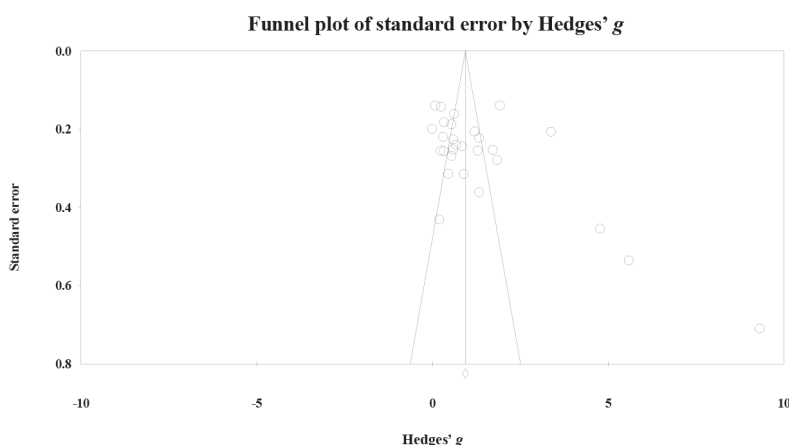
3. Results

3.1. Publication bias assessemnt

Figure 1 shows that the dots in this study are almost distributed symmetricaly around the vertical line. Additionally, although a few dots are outside the triangle

of the funnel plot, they are in the upper part of Figure 1 and not at the bottom. Borenstein et al. (2010) stated that a symmetric funnel plot implies that there is no publication bias. Therefore, it is concluded that the meta-analysis reliability of this study is not affected by publication bias.

FIGURE 1. Funnel plot of standard error by Hedges's g .



3.2. Effect of AI on students' learning achievement in CS education

Figure 2 presents the forest plot of the variation of effect size across the 28 included papers. The black square represents each paper's weighted effect size, where a larger square size implies a larger effect size. The arrow underneath each square (effect size) represents the confidence interval of the associated effect size. The overall mean effect size ($g = 1.364$) is presented in the last row of the forest plot. Overall, most papers had a positive effect size with different confidence intervals.

Table 2 shows that the meta-analysis yielded an overall effect size of $g = 1.36$, $p < .001$, indicating that AI had a very large

effect on students' learning achievement in CS education. Particularly, it is seen that all AI technologies had a significant positive effect, with some variation regarding the effect size; chatbots ($g = 1.35$, 95% CI = 0.77 to 1.94) and personalized learning systems ($g = 1.21$, 95% CI = 0.54 to 1.87) had a very large effect on students' learning achievement, while ITS ($g = 1.45$, 95% CI = 0.68 to 2.21) had a huge effect.

The I^2 statistic showed that 95.77% of variance resulted from between-study factors, implying that other variables might moderate the AI effect size on students' learning achievement. It is crucial, therefore, to conduct further analysis and investigate the potential moderating variables.

FIGURE 2. Forest plot of the Hedge's *g* estimates and the confidence intervals

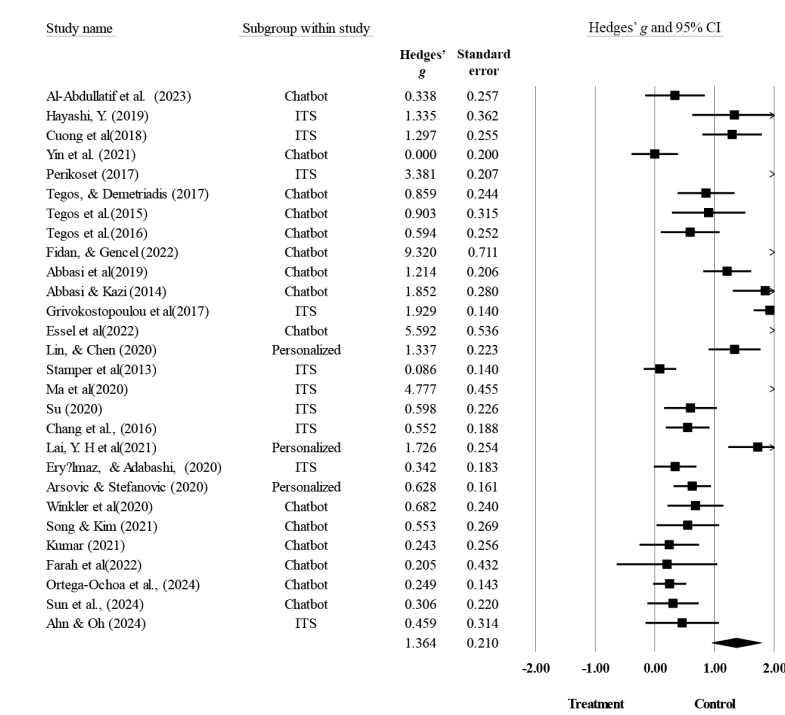


TABLE 2. Effect of AI on learning achievement in computer science education.

| Analysis | <i>n</i> | <i>g</i> | 95 % CI | <i>Z</i> | <i>p</i> | <i>I</i> ² | τ^2 | Effect size interpretation |
|----------------------|----------|----------|--------------|----------|----------|-----------------------|----------|----------------------------|
| Overall | 28 | 1.36 | [0.95, 1.78] | 6.50 | 0.001*** | 95.77 | 1.15 | Very large |
| Chatbot | 15 | 1.35 | [0.77, 1.94] | 4.54 | 0.001*** | 95.12 | 1.23 | Very large |
| ITS | 10 | 1.45 | [0.68, 2.21] | 3.71 | 0.001*** | 97.09 | 1.45 | Huge |
| Personalized systems | 3 | 1.21 | [0.54, 1.87] | 3.56 | 0.001*** | 87.18 | 0.30 | Very large |

Note: *n* = number of studies; *g* = Hedges' *g* effect size; CI = confidence interval; *Z* = *Z* value for Hedges' *g*; *p* = *p* values of Hedges' *g*; *I*² and τ^2 are measures of effect size variability; ****p* < .001.

3.3. Moderating variables of AI effect in CS education

To investigate the moderating variables of the AI effect, this section starts by first

investigating how the AI effect varies across various variables, including educational level (see Table 3), learning mode (see Table 4), AI intervention duration (see

Table 5), and geographical distribution of students (see Table 6). It then runs a meta-regression model taking into consideration all of the mentioned variables (see Table 7) to identify the moderating variables of the AI effect in CS education.

Table 3 presents the effect of AI on students' learning achievement across var-

ious educational levels. It is seen that AI was used mostly in higher education ($n = 26$) to teach about computer science subjects. The effect was very large ($g = 1.24$). On the other hand, one study for each educational level, namely primary and secondary education, found that the effects were very large ($g = 1.35$) and huge ($g = 4.77$), respectively.

TABLE 3. Effect of AI on learning achievement across various educational levels.

| Level of education | <i>n</i> | <i>g</i> | 95 % CI | <i>Z</i> | <i>p</i> | <i>I</i> ² | τ^2 | Effect size interpretation |
|--------------------|----------|----------|--------------|----------|----------|-----------------------|----------|----------------------------|
| PE | 1 | 1.35 | [0.63, 2.04] | 3.69 | 0.001*** | 0 | 0 | Very large |
| SE | 1 | 4.77 | [3.89, 5.67] | 10.50 | 0.001*** | 0 | 0 | Huge |
| HE | 26 | 1.24 | [0.83, 1.65] | 5.98 | 0.001*** | 95.58 | 1.04 | Very large |

Note: *n* = number of studies; *g* = Hedges' *g* effect size; CI = confidence interval; *Z* = *Z* value for Hedges' *g*; *p* = *p* values of Hedges' *g*; *I*² and τ^2 are measures of effect size variability; ****p* < .001; **p* < 0.5.

TABLE 4. Effect of AI on learning achievement in different learning modes.

| Modo de aprendizaje | <i>n</i> | <i>g</i> | 95 % CI | <i>Z</i> | <i>p</i> | <i>I</i> ² | τ^2 | Effect size interpretation |
|---------------------|----------|----------|--------------|----------|----------|-----------------------|----------|----------------------------|
| Blended | 4 | 0.69 | [0.23, 1.14] | 2.94 | 0.003** | 75.53 | 0.16 | Medium |
| Online | 24 | 1.49 | [1.02, 1.97] | 6.91 | 0.001*** | 96.31 | 1.31 | Huge |

Note: *n* = number of studies; *g* = Hedges' *g* effect size; CI = confidence interval; *Z* = *Z* value for Hedges' *g*; *p* = *p* values of Hedges' *g*; *I*² and τ^2 are measures of effect size variability; ****p* < .001.

The effect of AI in different learning modes on students' learning achievement was analyzed (see Table 4). It is evident that AI has a significant positive effect on learning achievement in all learning modes. Specifically, the application of AI in blended learning had a medium effect ($g = 0.69$), while the effect was very large in online learning ($g = 1.49$).

The effect of AI with different intervention durations on students' learning achievement was analyzed (see Table 5). It is seen that AI, regardless of the intervention duration, had a significant positive effect on learning achievement in CS education. Specifically, the effect of AI was huge ($g = 2.74$) when it was used between one week and one month.

However, this effect decreases (to be large) if AI is used for less or more than that period.

The effect of AI on students' learning achievement, based on their geographical distribution, was analyzed (see Table 6). It is seen that AI had a significant positive effect across all continents. Specifically, AI has a huge effect on the learning achievement of students in Africa ($g = 5.59$), while the effect was very large in Asia ($g = 1.10$) and Europe ($g = 1.59$). Fi-

nally, the effect was small in North America ($g = 0.26$).

To further investigate for possible covariance between the analyzed variables above, a meta-regression that includes all moderators was conducted. The obtained results revealed that the effect of AI on learning achievement in CS education significantly varies according to the AI intervention duration ($Q = 6.72, df = 2, p = 0.03$) and the geographical distribution of students ($Q = 14.47, df = 3, p = 0.001$).

TABLE 5. Effect of AI on learning achievement with different intervention durations.

| Duration | <i>n</i> | <i>g</i> | 95 % CI | <i>Z</i> | <i>p</i> | <i>I</i> ² | τ^2 | Effect size interpretation |
|---------------------------------|----------|----------|--------------|----------|----------|-----------------------|----------|----------------------------|
| Duration < 1 week | 2 | 0.72 | [0.33, 1.10] | 3.63 | 0.001*** | 0 | 0 | Large |
| 1 week ≤ duration < 1 month | 8 | 2.74 | [1.63, 3.84] | 4.86 | 0.001*** | 97.27 | 2.40 | Huge |
| 1 month ≤ duration < 1 semester | 18 | 0.85 | [0.52, 1.17] | 5.10 | 0.001*** | 90.53 | 0.44 | Large |

Note: *n* = number of studies; *g* = Hedges' *g* effect size; CI = confidence interval; *Z* = *Z* value for Hedges' *g*; *p* = *p* values of Hedges' *g*; *I*² and τ^2 are measures of effect size variability; ****p* < .001; **p* < .05.

TABLE 6. Effect of AI on learning achievement based the geographical distribution

| Continent | <i>n</i> | <i>g</i> | 95 % CI | <i>Z</i> | <i>p</i> | <i>I</i> ² | τ^2 | Effect size interpretation |
|---------------|----------|----------|---------------|----------|----------|-----------------------|----------|----------------------------|
| Africa | 1 | 5.59 | [4.54, 6.64] | 10.43 | 0.001*** | 0 | 0 | Huge |
| Asia | 14 | 1.10 | [0.65, 1.54] | 4.84 | 0.001*** | 91.38 | 0.64 | Very large |
| Europe | 11 | 1.59 | [0.83, 2.36] | 4.08 | 0.001*** | 97.26 | 1.58 | Very large |
| North America | 2 | 0.26 | [-0.18, 0.71] | 1.16 | 0.254 | 57.77 | 0.06 | Small |

Note: *n* = number of studies; *g* = Hedges' *g* effect size; CI = confidence interval; *Z* = *Z* value for Hedges' *g*; *p* = *p* values of Hedges' *g*; *I*² and τ^2 are measures of effect size variability; ****p* < .001; **p* < .05.

TABLE 7. Meta-regression results for the learning achievement of response from level of education, duration, geographical distribution, and learning mode.

| Model | Coefficient | Standard error | 95% lower | 95% upper | Z-value | 2 sided p value |
|-------------------------------------|---------------------------------|----------------|-----------|-----------|---------|-----------------|
| Intercept | 0.94 | 0.51 | -0.07 | 1.94 | 1.83 | 0.07 |
| Level of Education | 1 = PE | 1.41 | -3.89 | 1.65 | -0.79 | 0.43 |
| | 2 = SE | 1.56 | -0.14 | 5.99 | 1.87 | 0.06 |
| $Q^* = 4.92, df = 2, p = 0.09$ | | | | | | |
| Duration | 1 = Duration > 1 week | 1.00 | -2.40 | 1.51 | -0.45 | 0.66 |
| | 2 = 1 week ≤ duration < 1 month | 0.66 | 0.22 | 2.81 | 2.30 | 0.02* |
| $Q^* = 6.72, df = 2, p = 0.03^*$ | | | | | | |
| Geographical distribution | 1 = Africa | 1.42 | 2.48 | 8.04 | 3.71 | 0.001*** |
| | 2 = Europe | 0.62 | -0.97 | 1.48 | 0.41 | 0.68 |
| | 3 = North America | -0.62 | -2.56 | 1.32 | -0.63 | 0.53 |
| $Q^* = 14.47, df = 3, p = 0.001***$ | | | | | | |
| Learning mode | 1 = Blended | 0.75 | -1.82 | 1.12 | -0.47 | 0.64 |
| NA | | | | | | |

Note *** $p < .001$, * $p < .05$

4. Discussions

4.1. AI effect on students' learning achievement in CS education

This present study revealed that AI has a very large effect on learning achievement when it is used in computer science education compared to education generally, where several studies in the literature found that the effect was moderate (García-Martínez et al., 2023; Hwang, 2022; Wu & Yu, 2023). This could be because AI provides innovative tools and solutions that cater to the different advanced needs and competencies in computer science (Barnes et al., 2017). AI is also implemented based on a deep understanding of computer science pedagogy (Barnes et al., 2017), hence better contributing to learning achievement in CS education compared to other educational subjects/fields. Additionally, making full use of technology in education requires some competencies and a good background in understanding and using it. This requirement might be fulfilled by CS teachers who are considered to have good knowledge and expertise of AI (or technology generally) compared to other teachers from other educational fields (e.g., languages, science, etc.). Particularly, ITS, as an AI technology, had the highest achievement (huge) compared to other AI technologies, namely chatbots and personalized learning systems. This could be because ITSs provide instructions and interact with students at a finer level of granularity, allowing learning to be appropriate to students' different

needs and individual differences (Lin et al., 2023).

4.2. Moderating variables of the AI effect on students' learning achievement in CS education

This study revealed that educational level does not moderate the AI effect in CS education. While this present study did not focus during the search process on a specific educational level or on computer science as a curriculum (rather as educational subjects that are within CS), the obtained results revealed that AI is mostly used in higher education when teaching CS. This might be because computer science subjects are mostly taught in higher education, despite several attempts to teach some subjects like programming or AI at earlier educational levels (e. g., primary education or K-12). This could also be because students in higher education have the required competencies to use and work with AI compared to students in earlier educational levels. Particularly, the effect of AI on higher education students was very large in CS education. This might be because the integration of AI in higher education could easily adapt teaching to the various needs and profiles of learners (Verdú et al., 2017), provide individual and personalized feedback and improve assessments (Dever et al., 2020), and predict learning achievement (Çağataylı & Çelebi, 2022).

The results indicated that AI had a higher effect (a huge effect) when used to teach CS in online environments

than in blended environments (a medium effect). This could be explained by the fact that implementing collaborative intelligence (i.e., combining human intelligence and artificial intelligence) will be easier in online learning environments, as human-based and machine-based instructions can be designed, automated, and combined effectively in online learning models through the support of AI techniques and algorithms, as well as the collection and analysis of interaction (big) data of both teachers and students (i.e., learning analytics). Another reason might be because CS education requires the transfer of computational, mathematical, and algorithmic thinking (Pirker et al., 2014), which are better transferred through online instructions instead of using traditional instructions in classrooms. Although the learning mode did not moderate the effect of AI in CS education, the obtained results hold significance for educational developers and instructional designers, providing insights into potential integration of AI in future CS education and programs.

The results revealed that the AI had a huge effect on learning achievement when the intervention duration was between one week and one month. However, this effect decreased when AI was used for less or longer than this duration (between one week and one month). This might be because this duration is considered sufficient to promote students' familiarization and stimulate motivation for proper en-

gagement and performance (Merilampi et al., 2014), while a shorter time may be insufficient to produce the desired positive outcomes (Sung et al., 2016). It is further seen that the longer the AI intervention is, the lower the effect on learning achievement. This can be explained, as highlighted by several methodologists, by the fact that longer interventions might lead to poorer implementation fidelity, hence causing a shift from the initial experiment goals and achieving low effects (Mihalic, 2004; Wang et al., 2023). The decrease of the AI effect over longer periods might also be due to the novelty effect (Pisapia et al., 1993), where learners lose interest in a given technology as time passes. The intervention duration was found to significantly moderate the AI effect in CS education, revealing the need for systematic thinking about how long AI will be used and how it can achieve the desired educational objective(s) within a specific duration.

The results indicated that the AI effect on learning achievement in computer science education varies according to students' geographical distribution and also moderates the AI effect in CS education. This is because CS education has geographic particularities that should be considered (Francisco & De Oliveira, 2022). Particularly, the results revealed that AI has a huge effect on students learning achievement in Africa. However, this result cannot be generalized, as only one study was obtained in that sub-category (see

Table 6). On the other hand, the effect of AI on learning achievement was very large in Europe and Asia. This high effect could be attributed to the long and continuous efforts of the European Commission in developing AI literacy (Tangi, 2022) and promoting the safe and responsible implementation of AI (Jarota, 2023). This has led to clear policies, rich experiences, and organized educational systems that incorporate AI, resulting in better learning outcomes. Similarly, several Asian countries are now incorporating AI into their teaching and learning programs. This might have resulted in an improvement in learning outcomes. In this context, Nye (2015) highlighted that the community of AI in education is rapidly increasing and recognizing the importance of designing technologies globally.

5. Conclusions, implications and future directions

This study conducted a meta-analysis of 28 papers to investigate the AI effect on students' learning achievement in CS education. The obtained results revealed that AI has a very large effect on learning achievement. Particularly, the AI intervention duration and the geographical distribution of students are found to moderate this AI effect in CS education.

5.1. Implications

The findings of this study can contribute to the literature from various perspectives. From a theoretical

perspective, this study contributes to the ongoing debate about the effect of AI in education generally and in CS education specifically. It highlights what and how should be considered when implementing AI in CS education, hence contributing to developing frameworks and theories in this regard. Additionally, this study revealed a disparity in the AI effect within CS education across the geographical distribution of students. Therefore, more round-table discussions, policy making, and cross-country collaborations should be established to increase the effectiveness of using AI in CS education worldwide.

From a methodological perspective, the present study revealed that the AI effect on students' learning achievement in CS education might decrease over longer intervention periods. Therefore, researchers, pedagogues, and methodologists should rethink the AI intervention duration when designing their experiments or teaching approaches. In this context, Wang et al. (2023) highlighted that the effect of AI in education might be shaped by several confounding variables, including intervention duration, and more attention should be paid to these variables to accurately measure the AI effect and provide more generalizable results.

From a practical perspective, this present study revealed that AI can enhance students' learning achievement in CS education. It is therefore

important to raise awareness of the AI opportunities in education, thereby increasing its adoption and development in various CS disciplines. Additionally, the study revealed that developing a successful AI-powered educational technology is not an easy task because it should be treated as an ecosystem, where several confounding variables (education level, learning mode, intervention duration, geographical location, etc.) might impact the overall effectiveness of AI in education. This highlights the need for multidisciplinary collaboration (educators, computer scientists, data scientists, instructional designers, etc.) when developing AI systems in education generally and in CS education specifically. Finally, the findings of this study can contribute to achieving “quality education,” which is the fourth sustainable development goal (SDG) of the United Nations that several organizations and universities are working towards.

5.2. Limitations and future directions

Despite the reliability of this study, it still has some limitations that should be acknowledged and further researched. First, the findings of this study are limited by the search keywords and databases. Therefore, future research can complement the present research by considering more databases and search keywords. Additionally, this study did not investigate the effect of AI roles (e. g., tutor, teacher assistant, peer, instructor, etc.) as a potential moderator of the AI effect

in CS education. Such information is crucial to understanding how AI should be integrated into CS education to increase learning experiences and outcomes. It also did not consider CS disciplines when investigating the AI effect. Future studies can therefore address this by taking, for instance, the CS classification of the Association for Computing Machinery (ACM). Finally, this study did not delve much into the technical aspects of the implemented AI technologies. Future studies should build on this to investigate how different AI techniques and algorithms might moderate the effect of AI in CS education.

Author's contributions

Ahmed Tlili: Methodology; Software; Writing (original draft); Writing (review and editing).

References

References with an asterisk (*) indicate studies included in the analysis.

- AAAI. (2018). AAAI Launches “AI for K-12” Initiative in collaboration with the Computer Science Teachers Association (CSTA) and AI4All. <https://m.aaai.org/Pressroom/Releases/release-18-0515.pdf>
- *Abbasi, S., & Kazi, H. (2014). Measuring effectiveness of learning chatbot systems on student’s learning outcome and memory retention. *Asian Journal of Applied Science and Engineering*, 3 (2), 251-260.
- *Abbasi, S., Kazi, H., & Hussaini, N. N. (2019). Effect of chatbot systems on student’s learning outcomes. *Sylian*, 163 (10), 49-63.

- *Ahn, Y.-H., y Oh, E.-Y. (2024). Effects of the international training program for enhancing intelligent capabilities through blended learning on computational thinking, artificial intelligence competencies, and core competencies for the future society in graduate students. *Applied Sciences*, 14 (3). <https://doi.org/10.3390/app14030991>
- *Al-Abdullatif, A. M., Al-Dokhny, A. A., & Drwish, A. M. (2023). Implementing the Bashayer chatbot in Saudi higher education: Measuring the influence on students' motivation and learning strategies. *Frontiers in psychology*, 14, 1129070. <https://doi.org/10.3389/fpsyg.2023.1129070>
- *Arsovic, B., & Stefanovic, N. (2020). E-learning based on the adaptive learning model: Case study in Serbia. *Sādhana*, 45 (1), 266. <https://doi.org/10.1007/s12046-020-01499-8>
- Barnes, T., Boyer, K., Hsiao, S. I. H., Le, N. T., & Sosnovsky, S. (2017). Preface for the special issue on AI-supported education in computer science. *International Journal of Artificial Intelligence in Education*, 27, 1-4. <https://doi.org/10.1007/s40593-016-0123-y>
- Borenstein, M. (2022). Comprehensive meta-analysis software. In M. Egger, J. P. T. Higgins, & G. D. Smith (Eds.), *Systematic reviews in health research: Meta-analysis in context* (pp. 535-548). BMJ Books. <https://doi.org/10.1002/9781119099369.ch27>
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, 1 (2), 97-111. <https://doi.org/10.1002/jrsm.12>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). John Wiley & Sons.
- Çağataylı, M., & Çelebi, E. (2022). Estimating academic success in higher education using big five personality traits, a machine learning approach. *Arab Journal Scientific Engineering*, 47, 1289-1298. <https://doi.org/10.1007/s13369-021-05873-4>
- *Chang, Y. H., Chen, Y. Y., Chen, N. S., Lu, Y. T., & Fang, R. J. (2016). Yet another adaptive learning management system based on Felder and Silverman's learning styles and Mashup. *Eurasia Journal of Mathematics, Science and Technology Education*, 12 (5), 1273-1285. <https://doi.org/10.12973/eurasia.2016.1512a>
- Crow, T., Luxton-Reilly, A., & Wuensche, B. (2018). Intelligent tutoring systems for programming education: A systematic review. In *Proceedings of the 20th Australasian Computing Education Conference* (pp. 53-62). Association for Computing Machinery.
- *Cuong, N. D. H., Arch-Int, N., & Arch-Int, S. (2018). FUSE: A fuzzy-semantic framework for personalizing learning recommendations. *International Journal of Information Technology & Decision Making*, 17 (4), 1173-1201. <https://doi.org/10.1142/S0219622018500220>
- Dever, D. A., Azevedo, R., Cloude, E. B., & Wiedbusch, M. (2020). The impact of autonomy and types of informational text presentations in game-based environments on learning: Converging multi-channel processes data and learning outcomes. *International Journal of Artificial Intelligence in Education*, 30 (4), 581-615. <https://doi.org/10.1007/s40593-020-00215-1>
- *Eryilmaz, M., & Adabashi, A. (2020). Development of an intelligent tutoring system using Bayesian networks and fuzzy logic for a higher student academic performance. *Applied Sciences*, 10 (19), 6638. <https://doi.org/10.3390/app10196638>

- *Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19 (1), 1-19. <https://doi.org/10.1186/s41239-022-00362-6>
- *Farah, J. C., Spaenlehauer, B., Sharma, V., Rodríguez-Triana, M. J., Ingram, S., & Gillet, D. (2022). Impersonating chatbots in a code review exercise to teach software engineering best practices. In *2022 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1634-1642). IEEE.
- *Fidan, M., & Gencel, N. (2022). Supporting the instructional videos with chatbot and peer feedback mechanisms in online learning: The effects on learning performance and intrinsic motivation. *Journal of Educational Computing Research*, 60 (7), 1716-1741. <https://doi.org/10.1177/07356331221077901>
- Francisco, R. E., & de Oliveira, F. (2022). Intelligent tutoring system for computer science education and the use of artificial intelligence: A literature review. *CSEDU*, (1), 338-345.
- García-Martínez, I., Fernández-Batanero, J., Fernández-Cerero, J., & León, S. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: Systematic review and meta-analysis. *Journal of New Approaches in Educational Research*, 12 (1), 171-197. <https://doi.org/10.7821/naer.2023.1.1240>
- Gerdes, A., Heeren, B., Jeurig, J., & Van Binsbergen, L. T. (2017). Ask-Elle: An adaptable programming tutor for Haskell giving automated feedback. *International Journal of Artificial Intelligence in Education*, 27, 65-100. <https://doi.org/10.1007/s40593-015-0080-x>
- Goel, A. K., & Joyner, D. A. (2017). Using AI to teach AI: Lessons from an online AI class. *Ai Magazine*, 38 (2), 48-59. <https://doi.org/10.1609/aimag.v38i2.2732>
- *Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2017). An educational system for learning search algorithms and automatically assessing student performance. *International Journal of Artificial Intelligence in Education*, 27, 207-240. <https://doi.org/10.1007/s40593-016-0116-x>
- *Hayashi, Y. (2019). Multiple pedagogical conversational agents to support learner-learner collaborative learning: Effects of splitting suggestion types. *Cognitive Systems Research*, 54, 246-257. <https://doi.org/10.1016/j.cogsys.2018.04.005>
- Hedges, L. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6 (2), 107-128. <https://doi.org/10.3102/10769986006002107>
- Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.
- Hwang, S. (2022). Examining the effects of artificial intelligence on elementary students' mathematics achievement: A meta-analysis. *Sustainability*, 14 (20), 1-20. <https://doi.org/10.3390/su142013185>
- Jarota, M., (2023). Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective. *Computer Law & Security Review*, 49, 1-14. <https://doi.org/10.1016/j.clsr.2023.105825>
- Kim, J., & Lee, SS. (2023). Are two heads better than one? The effect of student-AI collaboration on students' learning task performance. *TechTrends*, 67, 365-375. <https://doi.org/10.1007/s11528-022-00788-9>
- Kitchenham, B. A., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele University and Durham University.

- *Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. *International journal of educational technology in higher education*, 18 (1), 1-28. <https://doi.org/10.1186/s41239-021-00302-w>
- *Lai, Y. H., Chen, S. Y., Lai, C. F., Chang, Y. C., & Su, Y. S. (2021). Study on enhancing AIoT computational thinking skills by plot image-based VR. *Interactive Learning Environments*, 29 (3), 482-495. <https://doi.org/10.1080/10494820.2019.1580750>
- *Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational technology research and development*, 70 (5), 1843-1865. <https://doi.org/10.1007/s11423-022-10142-8>
- Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of General Internal Medicine*, 33 (8), 1260-1267. <https://doi.org/10.1007/s11606-018-4425-7>
- Lin, R., Zhang, Q., Xi, L., & Chu, J. (2022). Exploring the effectiveness and moderators of artificial intelligence in the classroom: A meta-analysis. In J. Yang, D. Liu, Kinshuk, A. Tlili, M. Chang, E. Popescu, D. Burgos, & Z. Altınay (Eds.), *Resilience and future of smart learning. Proceedings of 2022 International Conference on Smart Learning Environments*. Springer. https://doi.org/10.1007/978-981-19-5967-7_7
- Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments*, 10 (1), 41. <https://doi.org/10.1186/s40561-023-00260-y>
- *Ma, Z. H., Hwang, W. Y., & Shih, T. K. (2020). Effects of a peer tutor recommender system (PTRS) with machine learning and automated assessment on vocational high school students' computer application operating skills. *Journal of Computers in Education*, 7, 435-462. <https://doi.org/10.1007/s40692-020-00162-9>
- Mihalic, S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth*, 4 (4), 83-105.
- Merilampi, S., Sirkka, A., Leino, M., Koivisto, A. and Finn, E. (2014). Cognitive mobile games for memory impaired older adults. *Journal of Assistive Technologies*, 8 (4), 207-223. <https://doi.org/10.1108/JAT-12-2013-0033>
- Nesbit, J. C., Adesope, O. O., Liu, Q., & Ma, W. (2014). How effective are intelligent tutoring systems in computer science education? In *The 14th IEEE International Conference on Advanced Learning Technologies (ICALT 2014). Advanced Technologies for Supporting Open Access to Formal and Informal Learning* (pp. 99-103). IEEE.
- Nesbit, J., Liu, L., Liu, Q., and Adesope, O. O. (2015). Work in progress: Intelligent tutoring systems in computer science and software engineering education. In *2015 ASEE Annual Conference and Exposition Proceedings* (pp. 26.1754.1-26.1754.12). American Society for Engineering Education.
- Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, 25, 177-203. <https://doi.org/10.1007/s40593-014-0028-6>
- Onesi-Ozigagun, O., Ololade, Y. J., Eyo-Udo, N. L., & Ogundipe, D. O. (2024). Revolutionizing education through AI: A comprehensive review of enhancing learning experiences. *International Journal of Applied Research in Social Sciences*, 6 (4), 589-607. <https://doi.org/10.51594/ijarss.v6i4.1011>

- *Ortega-Ochoa, E., Quiroga, J., Arguedas, M., Daradoumis, T., & Marquès, J. M. (2024). The effectiveness of empathic chatbot feedback for developing computer competencies, motivation, self-regulation, and metacognitive reasoning in online higher education. *Internet of Things*, 25, 101101. <https://doi.org/10.1016/j.iot.2024.101101>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, M. A., Stewart, L. A., Thomas, J., Tricco, A. C., ..., y Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International journal of surgery*, 88, 105906.
- *Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2017). Assistance and feedback mechanism in an intelligent tutoring system for teaching conversion of natural language into logic. *International Journal of Artificial Intelligence in Education*, 27, 475-514. <https://doi.org/10.1007/s40593-017-0139-y>
- Pirker, J., Riffnaller-Schiefer, M., & Gütl, C. (2014). Motivational active learning: Engaging university students in computer science education. In *Proceedings of the 2014 Conference on Innovation y Technology in Computer Science Education* (pp. 297-302). Association for Computing Machinery.
- Pisapia, J., Schlesinger, J., & Parks, A. (1993). *Learning technologies in the classroom: Review of the literature*. Metropolitan Educational Research Consortium.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86 (3), 638-641. <https://doi.org/10.1037//0033-2909.86.3.638>
- Shaw, M. (1985). The nature of computer science. In M. Shaw (Ed.), *The Carnegie-Mellon curriculum for undergraduate computer science*. Springer.
- Silva, D. B., Aguiar, R. d. L., Dvconlo, D. S., & Silla, C. N. (2019). Recent studies about teaching algorithms (CS1) and data structures (CS2) for computer science students. In *2019 IEEE Frontiers in Education Conference (FIE)* (pp. 1-8). IEEE.
- Smith, M. (2016, 30 January). Computer science for all [blog]. *The White House - President Barack Obama*. <https://www.whitehouse.gov/blog/2016/01/30/computer-science-all>
- *Song, D., & Kim, D. (2021). Effects of self-regulation scaffolding on online participation and learning outcomes. *Journal of Research on Technology in Education*, 53 (3), 249-263. <https://doi.org/10.1080/15391523.2020.1767525>
- *Stamper, J., Eagle, M., Barnes, T., & Croy, M. (2013). Experimental evaluation of automatic hint generation for a logic tutor. *International Journal of Artificial Intelligence in Education*, 22 (1-2), 3-17. https://doi.org/10.1007/978-3-642-21869-9_45
- *Su, J. M. (2020). A rule-based self-regulated learning assistance scheme to facilitate personalized learning with adaptive scaffoldings: A case study for learning computer software. *Computer Applications in Engineering Education*, 28 (3), 536-555. <https://doi.org/10.1002/cae.22222>
- *Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, 21, 14. <https://doi.org/10.1186/s41239-024-00446-5>

- Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, 94, 252-275. <https://doi.org/10.1016/j.compedu.2015.11.008>
- Tangi, L., van Noordt, C., Combetto, M., Gattwinkel D., & Pignatelli F. (2022). *AI watch. European landscape on the use of artificial intelligence by the public sector*. Publications Office of the European Union. <https://doi.org/10.2760/39336>
- *Tegos, S., & Demetriadis, S. (2017). Conversational agents improve peer learning through building on prior knowledge. *Journal of Educational Technology & Society*, 20 (1), 99-111.
- *Tegos, S., Demetriadis, S., & Karakostas, A. (2015). Promoting academically productive talk with conversational agent interventions in collaborative learning settings. *Computers & Education*, 87, 309-325. <https://doi.org/10.1016/j.compedu.2015.07.014>
- *Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: A comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, 11 (4), 417-440. <https://doi.org/10.1007/s11412-016-9246-2>
- Tlili, A., Essalmi, F., & Jemni, M. (2015). An educational game for teaching computer architecture: Evaluation using learning analytics. In *2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA)* (pp. 1-6). IEEE.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10 (1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Touretzky, D., Gardner-McCune, C., & Seehorn, D. (2023). Machine learning and the five big ideas in AI. *International Journal of Artificial Intelligence in Education*, 33 (2), 233-266. <https://doi.org/10.1007/s40593-022-00314-1>
- Verdú, E., Regueras, L. M., Gal, E., De Castro, J. P., Verdú, M. J., & Kohen-Vacs, D. (2017). Integration of an intelligent tutoring system in a course of computer network design. *Educational Technology Research and Development*, 65, 653-677. <https://doi.org/10.1007/s11423-016-9503-0>
- Waraich, A. (2004). Using narrative as a motivating device to teach binary arithmetic and logic gates. *ACM SIGCSE Bulletin*, 36 (3), 97-101. <https://doi.org/10.1145/1026487.1008024>
- Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and information technologies*, 28 (7), 9113-9148. <https://doi.org/10.1007/s10639-022-11555-x>
- *Winkler, R., Hobert, S., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020, April). Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). Association for Computing Machinery.
- Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, 55 (1), 10-33. <https://doi.org/10.1111/bjet.13334>
- *Yin, J., Goh, T. T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, 59 (1), 154-177. <https://doi.org/10.1177/0735633120952067>

- Zhang, X., Tlili, A., Shubeck, K., Hu, X., Huang, R., & Zhu, L. (2021). Teachers' adoption of an open and interactive e-book for teaching K-12 students artificial intelligence: A mixed methods inquiry. *Smart Learning Environments*, 8, 1-20. <https://doi.org/10.1186/s40561-021-00176-5>
- Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2023). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, 31 (9), 5650-5664. <https://doi.org/10.1080/10494820.2021.2015693>

Author's biography

Ahmed Tlili. Associate professor of Educational Technology at Beijing Normal University, China, and adjunct associate professor at An-Najah University, Palestine. He is the co-director of the OER Lab at the Smart Learning Institute of Beijing Normal University (SLIBNU), China. He serves as the editor of Springer Series *Future Education and Learning Spaces*, and the

deputy-editor-in-chief of *Smart Learning Environments* journal. He is also the associate editor of *IEEE Bulletin of the Technical Committee on Learning Technology*, and the journal of *e-Learning and Knowledge Society*. Professor Tlili is also an expert at the Arab League Educational, Cultural and Scientific Organization (ALECSO). He has edited several special issues in several journals. He has also published several books, as well as academic papers in international referred journals and conferences. He has been awarded the Martin Wolpers 2021 Prize by the Research Institute for Innovation and Technology in Education (UNIR iTED) in recognition of excellence in research, education and significant impact on society. He also has been awarded the IEEE TCLT Early Career Researcher Award in Learning Technologies for 2020.



<https://orcid.org/0000-0003-1449-7751>

