DOI: 10.5961/higheredusci.1747885

Artificial Intelligence Threatens Critical Thinking in Education Systems

Yapay Zekâ Eğitim Sistemlerinde Eleştirel Düşünmeyi Tehdit Ediyor

Mahmut ÖZER, Hande TANBERKAN, Matjaz PERC

ABSTRACT

We examine how the increasing use of artificial intelligence (AI) in education—through tools that generate and summarize text, translate languages, and produce visual content—impacts students' critical thinking. While these technologies enhance personalized learning, broaden assessment strategies, and support data-driven policy decisions, we argue that their integration into the learning process carries unintended cognitive consequences. Specifically, we show that when students offload key tasks to AI systems, their cognitive load decreases in ways that weaken memory retention and reduce active engagement with content. This shift fosters a pattern of overreliance, as students increasingly depend on AI to perform intellectual tasks in their place. As a result, their ability to think critically, question information, and evaluate sources diminishes over time. We highlight this emerging dependency as a medium- to long-term threat to critical thinking and call for a more careful evaluation of how generative AI is used in education—not only in terms of its benefits, but also its influence on core cognitive processes. Finally, we propose targeted strategies to mitigate these effects and preserve students' critical capacities in AI-rich learning environments.

Keywords: Artificial intelligence, Education system, Critical thinking, Cognitive load, Memory

ÖZ

Bu çalışmada eğitimde metin üreten ve özetleyen, dilleri çeviren ve görsel içerik oluşturan araçlar yoluyla yapay zekânın giderek artan kullanımının öğrencilerin eleştirel düşünme becerileri üzerindeki etkisini inceliyoruz. Bu teknolojiler kişiselleştirilmiş öğrenmeyi geliştirse, değerlendirme stratejilerini çeşitlendirse ve veriye dayalı politika kararlarını desteklese de, öğrenme sürecine entegrasyonlarının istenmeyen bilişsel sonuçlar doğurduğunu savunuyoruz. Özellikle, öğrencilerin temel görevleri yapay zekâ sistemlerine devrettiğinde, bilişsel yükleri öyle bir şekilde azalıyor ki bu durum hem hafıza kalıcılığını zayıflatıyor hem de içeriğe yönelik aktif katılımı azaltıyor. Bu değişim, öğrencilerin entelektüel görevleri kendi yerlerine yapay zekânın gerçekleştirmesine giderek daha fazla güvenmesiyle bir aşırı

Özer M., Tanberkan H., & Perc M., (2025). Artificial intelligence threatens critical thinking in education systems. *Journal of Higher Education and Science/Yükseköğretim ve Bilim Dergisi, 15*(2), 157-164. https://doi.org/10.5961/higheredusci.1747885

Mahmut ÖZER
ORCID ID: 0000-0001-8722-8670

Grand National Assembly of Türkiye, Ankara, Türkiye
Türkiye Büyük Millet Meclisi, Ankara, Türkiye

Hande TANBERKAN (✉)
ORCID ID: 0000-0001-7142-5397

Baskent University, Faculty of Education, Department of Educational Sciences, Ankara, Türkiye
Başkent Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Ankara, Türkiye
handetanberkan@gmail.com

Matjaz PERC
ORCID ID: 0000-0002-3087-541X

University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor, Slovenia
Maribor Üniversitesi, Doğa Bilimleri ve Matematik Fakültesi, Maribor, Slovenya

Received/*Geliş Tarihi* : 22.07.2025
Accepted/*Kabul Tarihi* : 18.08.2025

bağımlılık örüntüsünü teşvik ediyor. Sonuç olarak, eleştirel düşünme, bilgiyi sorgulama ve kaynakları değerlendirme becerileri zamanla zayıflıyor. Bu gelişen bağımlılığı, eleştirel düşünme açısından orta ve uzun vadeli bir tehdit olarak vurguluyor ve üretken yapay zekânın eğitimde nasıl kullanıldığının yalnızca faydaları açısından değil, temel bilişsel süreçler üzerindeki etkileri bakımından da dikkatlice değerlendirilmesi gerektiğini savunuyoruz. Son olarak, bu etkileri azaltmaya ve yapay zekâ yoğun öğrenme ortamlarında öğrencilerin eleştirel kapasitesini korumaya yönelik stratejiler öneriyoruz.

Anahtar Sözcükler: Yapay zekâ, Eğitim sistemi, Eleştirel düşünme, Bilişsel yük, Hafıza

INTRODUCTION

Artificial intelligence continues to transform every domain, from education to healthcare, from the financial sector to service industries, and from the pharmaceutical industry to defense (Ilikhan et al., 2024; 2025; Özer, 2024a; 2024b; 2024c; Perc et al., 2019; Suleyman, 2023). In particular, the widespread adoption of generative AI applications such as ChatGPT and DeepSeek has significantly accelerated both its diffusion and the ensuing transformation. Unlike previous general-purpose technological disruptions, this new breakthrough has rapidly permeated all aspects of life. Today, we are witnessing the construction of a powerful AI ecosystem.

As the proliferation of artificial intelligence increases, significant disruptions have begun to emerge in the labor market. Expectations surrounding many occupations and job positions are changing rapidly. While the skill requirements for these roles continue to evolve, concerns about AI's negative impact on employment are also growing (Özer et al., 2024a; 2024b). As with previous general-purpose technological transformations, the initial expectation was that the number of new jobs created would roughly match the number of jobs displaced. However, the rapid advancement of AI applications has cast this expectation in a more pessimistic light (Özer & Perc, 2024). Today, with the rapid acceleration of automation through AI, many existing occupations and job positions are becoming obsolete, while the skill demands for newly emerging roles are rising significantly. Moreover, it is predicted that even these newly created roles will eventually be filled by AI (Suleyman, 2023). Consequently, large segments of the population are being forced to acquire new skills in order to maintain their current jobs, while others—due to age or background—are gradually losing their jobs and are being pushed toward lower-skilled and therefore lower-paid employment. The massive wave brought by AI now stands as the most significant threat to the increasingly eroded middle class in many countries (Özer, 2024d).

Given these circumstances, it is the education systems that are naturally the most affected by this wave (Özer, 2024a). Today, education systems face a two-dimensional challenge. First, due to AI's impact on the labor market and the resulting shifts in the skills demanded by occupations, education systems are now required to train human capital in alignment with these new skill expectations. On the one hand, they are being pressured to rapidly implement changes that align with the current transformation; on the other hand, they must also respond to the far more complex challenge of preparing individuals for jobs that do not yet exist but may emerge in the future. Second, just as AI technologies are transforming business processes across every sector, education systems now bear an additional responsibility: figuring out how to integrate these technologies into learning environments while also mitigating their potential risks.

As with previous general-purpose technological transformations, the benefits brought by the AI revolution have taken center stage. Expectations have steadily grown that these applications will transform educational processes for students, teachers, and administrators alike—particularly that generative AI tools will enrich learning environments and significantly expand the availability of personalized education options. Consequently, this sense of anticipation has rapidly spread to encompass all stakeholders in education, including students, teachers, school administrators, and families.

Notably, students across all levels of the education system have begun actively using these tools. In fact, in many countries, it is known that students are using AI applications much more frequently than teachers. In some contexts, this widespread use initially raised concerns among education administrators, leading to temporary bans on AI tools in schools and on campuses. However, given the evident benefits of these technologies and the difficulty of resisting such a profound transformation, the focus has increasingly shifted from prohibition to developing frameworks that enhance their advantages while mitigating their associated risks.

The capabilities of generative AI applications—such as text generation, text summarization, academic writing, cross-lingual translation, and the creation of graphics, images, and videos—have significantly supported students in fulfilling their academic responsibilities (Özer, 2024a). In particular, the potential of these tools to enhance personalized learning options quickly captured the attention not only of students, but also of teachers and education administrators (Ambele et al., 2022; Grassini, 2023; Kasneci et al., 2023). Actively leveraging personalized learning opportunities presents a critical chance to address one of the most persistent challenges in education systems: reducing achievement gaps between schools and strengthening educational equity. Furthermore, generative AI offers significant opportunities to enhance the quality of foreign language, music, and fine arts education as well (Gangadharbatla, 2022; Özer, 2024a; Wang et al., 2023; Zylinska, 2023).

AI applications also hold the potential to significantly transform the traditional roles of teachers. In particular, the additional materials generated through these tools have made it increasingly feasible to enrich learning environments and ensure the active participation of all students (Atlas, 2023). Moreover, providing targeted support to students with learning gaps within a classroom has become much easier thanks to personalized learning aids. As a result, teachers are now expected to move beyond their conventional roles and adopt new responsibilities—much like conductors—guiding each student to contribute at an equal level to the educational "orchestra" (Rudolph et al., 2023; Özer, 2024a). By using intelligent tutoring systems (ITS), teachers can now generate individualized instructional content tailored to help each student engage at a comparable level (Rudolph et al., 2023; Zawacki-Richter et al., 2019). Furthermore, these tools have also simplified the creation of alternative assessment formats, enabled rapid evaluation, and facilitated timely feedback for students (Grassini, 2023; Khan et al., 2023; Suna & Özer, 2025; Tanberkan et al., 2024; Wang et al., 2023).

On the other hand, the potential of AI applications to enrich the systemic perspectives of education administrators—from the school level to the national system level—has quickly become apparent (Özer, 2024a). In this context, such tools can clarify the steps needed to enhance educational quality at the school level, from guiding students in course selection to monitoring teacher and classroom performance; at the broader system level, they can also make significant contributions by enabling the development of evidence-based education policies through the effective use of big data (Chiu, 2023; Tsai et al., 2020; Villegas-Ch et al., 2020).

Although the enthusiasm surrounding the benefits of AI applications has so far overshadowed concerns, the risks associated with these technologies have increasingly begun to surface and spark debate (Suleyman, 2023). Chief among these risks—just as in other domains—is data security. When the data generated through the widespread use of AI in education systems is not adequately protected, there is a real danger that individuals may become vulnerable, manipulable commercial objects, especially when this educational data is combined with personal data collected through other applications in everyday life (O'Neil, 2016; Özer, 2025). Another major risk lies in the reproduction of biases—either stemming from the assumptions embedded in algorithmic design or from the vast training data itself, which may contain biases based on religion, culture, race, and other factors (Obermeyer et al., 2019; Özer et al., 2024a; Ulnicane & Aden, 2023). Similarly, it has long been known that AI can experience "hallucinations," generating information that appears coherent and credible but is actually incorrect (Özer, 2024c). The ability of AI to produce biased or inaccurate content stands out as one of the most critical risks for the education system.

On the other hand, recent studies indicating that artificial intelligence weakens critical thinking within education systems have raised growing concerns about what may be its most detrimental long-term impact. For this reason, the present study re-evaluates the threats posed by the integration of AI applications into education with a specific focus on their effects on critical thinking, and offers recommendations aimed at mitigating these risks.

Biased Content Is Becoming More Pervasive

AI applications operate based on algorithms that learn from real-world data to optimize and generate content. The assumptions made during algorithm design, as well as the biases embedded in training datasets, directly shape the nature of the content produced by AI. Consequently, biases related to religion, culture, race, gender, and other identity markers—whether stemming from algorithmic assumptions or training data—pose a significant risk of becoming more widespread as these applications rapidly permeate everyday life. Excessive trust in AI-generated content facilitates the unchecked spread of these biases and contributes to the deepening of social inequalities.

Assumptions made during algorithm development often contain inherent biases (O'Neil, 2016; Özer, 2025). When these assumptions are biased, the content generated by the algorithm inevitably reproduces these biases. A notable example of biased assumptions in algorithms can be found in a widely used AI application in the United States designed to identify individuals in need of advanced healthcare. A study revealed that this application used healthcare expenditures as a proxy to determine the need for advanced medical services (Obermeyer et al., 2019). According to this assumption, individuals who spend more on healthcare are presumed to have a greater need for advanced services. However, this overlooks structural inequalities in access to healthcare and differences in socioeconomic status. As a result of this flawed assumption, white individuals—who tend to spend more on healthcare—were more likely to be identified as needing advanced care. In contrast, Black Americans, who may have had equal or even greater healthcare needs but lacked access to services or the financial means to incur such expenditures, were underrepresented. Thus, the AI system, based on a biased assumption, exacerbated existing disadvantages among already marginalized individuals.

Algorithmic bias is also present in assumptions based on academic achievement. It has long been established that there is a strong correlation between academic success and a student's socioeconomic status (Özer & Suna, 2022; Suna et al., 2020; Suna & Özer, 2021; Suna et al., 2021; Suna & Özer, 2022; Suna & Özer, 2024). Therefore, when academic transitions within the education system or admissions to different institutions rely solely on academic performance indicators—such as GPA or standardized test scores like the SAT or ACT—the algorithm's output implicitly ends up ranking applicants according to their socioeconomic background (Sackett et al., 2009). We know that as socioeconomic status (SES) increases, students tend to perform better on high-stakes exams such as the SAT or ACT (Özer et al., 2024a). A recent and well-known example occurred in the United Kingdom during the COVID-19 pandemic, when an algorithm used to assign student grades sparked public outrage. The algorithm awarded higher grades to students from socioeconomically advantaged backgrounds while assigning lower grades to disadvantaged students, leading to widespread societal backlash (Heaton et al., 2023; Idowu, 2024).

Similar biases are also reproduced from the data on which AI applications are trained. Since AI algorithms learn from training datasets, and these datasets function as a form of memory, any biases embedded in that memory inevitably permeate the content generated by such systems. Unlike algorithmic design flaws, the biases found in training datasets are not mere technical errors; rather, they reflect the underlying power structures, socioeconomic conditions, and political systems of society (Ulnicane & Aden, 2023). For example, numerous studies in the U.S. context have revealed that AI systems used to determine sentencing durations based on recidivism risk exhibit systemic discrimination against Black Americans, recommending harsher penalties compared to those suggested for white defendants (Angwin et al., 2016; Dressel & Farid, 2018). Similar patterns have been observed in AI applications used to guide daily police patrol routes (Lum & Isaac, 2016). Because crime records often originate from areas predominantly inhabited by low-income individuals, Black communities, and minorities, algorithms trained on such data tend to disproportionately direct police patrols to these same areas—neglecting other regions with similar potential for crime. In this way, the advantages and disadvantages embedded in real-world data are reproduced by AI systems, exacerbating the very inequalities they reflect. In other words, such systems deepen existing social disparities.

A similar phenomenon can be observed in the texts generated by AI. Generative AI systems learn from vast datasets and produce new texts based on that learning; in this sense, the existing data serves as a kind of memory embedded in the newly generated content. This is particularly evident in the context of Orientalism. Orientalism, by establishing cultural hegemony, denies the non-Western world the agency to represent itself and confines it within representations produced by the West (Said, 1979). Given that a vast body of Orientalist literature has been built up over the past two centuries, generative AI systems tend to reproduce this same historical memory when generating content about the non-Western world. AI applications often fail to produce culturally accurate representations and instead tend to remove content from its original context or assemble it through a patchwork approach. This indicates that generative AI continues to reflect an Orientalist tone, wherein Western and white perspectives dominate as cultural defaults (Ghosh et al., 2024; Qadri et al., 2023). In this way, AI technologies risk reinforcing and disseminating Orientalist discourses.

In short, the assumptions made during the development of AI applications, along with the data on which they are trained, are directly reflected in the outputs they generate. For this reason, users must approach AI-generated content with caution, actively identify and filter out biases, and, most importantly, recognize that not all information produced by AI is accurate. It is essential to foster a habit of critically evaluating AI-generated content. However, the continual reinforcement of overreliance on these tools weakens critical thinking and contributes to the unchecked spread of biased information.

AI Hallucinates

In this context, another major risk to critical thinking posed by the use of AI in education is its ability to generate content that is unrelated to the given prompt or factually incorrect, even though it may appear coherent and internally consistent (Özer, 2024b). This phenomenon is referred to as AI hallucination (Ji et al., 2023). For instance, studies have shown that many of the references produced by AI during the drafting of academic papers are fabricated and do not exist in reality (Athaluri et al., 2023). It is argued that this hallucinatory behavior arises when the AI loses its connection to the original training dataset (Berberette et al., 2024). Most critically, once the AI begins generating inaccurate content based on such behavior, it tends to maintain this consistency—continuing to hallucinate in a seemingly logical manner. This escalating effect is known as the "snowball effect of hallucination" (Zhang & Press et al., 2023).

It has been suggested that AI systems are particularly prone to hallucination when faced with conflicting information within their training data, leading to a kind of cognitive tension during response generation (Özer, 2024c). Regardless of the source, this behavior indicates that AI-generated content may contain inaccurate or false information. Therefore, as with the biased content discussed above, users should avoid accepting AI-generated information at face value. Instead, it is essential to actively engage critical thinking skills—filtering, evaluating, and, if necessary, discarding information—to ensure accuracy and reliability.

Increases Cognitive Cost

Recent studies suggest that beyond the risks mentioned above, artificial intelligence poses serious threats specifically to critical thinking. For example, recent research conducted with university students has found that a significant portion of student laziness is linked to the use of AI applications (Ahmad et al., 2023). Moreover, a large number of students acknowledged that due to these tools, they focus less on solving problems independently, which in turn has weakened their critical thinking capacity (Mohammadkarimi & Omar, 2025). Both studies indicate that AI applications are taking over a substantial part of students' responsibilities, rendering them more passive in the learning process and ultimately diminishing their ability to think critically.

Further support for the claim that AI applications weaken critical thinking comes from a recent study. This study examined the relationship between the use of AI tools and critical thinking skills through the lens of the phenomenon of cognitive offloading. The findings reveal that AI tools indeed diminish critical thinking capabilities, with cognitive offloading playing a pivotal role in this decline (Gerlich, 2025). Excessive reliance on these tools increases cognitive offloading and reduces individuals' active engagement in cognitive processes. Notably, younger participants—who exhibited higher levels of dependency on AI tools compared to older participants—scored lower in critical thinking assessments. This serves as a significant early warning about long-term challenges we may face. In addition, the study found that participants with higher levels of education still demonstrated stronger critical thinking skills, highlighting the critical role of education in mitigating this negative effect.

The findings above indirectly highlight the critical importance of how much external technological tools used during education encourage students' active participation in learning processes. As active engagement decreases, cognitive offloading increases, and as a result, students' critical thinking skills fail to develop through new experiential learning. A recent study examining the relationship between active engagement in cognitive processes and the nature of external tools also points to this connection (Stadler et al., 2024). In this study, which measured the impact of different external tools on university students' ability to develop arguments on a given topic, two tools were used: large language models (LLMs) and traditional search engines. While LLMs directly provided compiled information to students, traditional search engines only displayed relevant links requiring students to evaluate, analyze, and incorporate the information from those sources into their arguments—thereby fostering deeper cognitive involvement.

The findings of the study indicate that students who used LLMs experienced lower cognitive load across all three dimensions—namely intrinsic, extraneous, and germane cognitive load—compared to those who used traditional search engines. This suggests that, by their very nature, LLMs assume a substantial portion of the user's cognitive load, thereby reducing the level of active cognitive engagement in the learning process. This effect is also observable in the diversity of valid arguments students developed. The study found that students using traditional search engines produced a greater variety of valid arguments, whereas argument diversity was notably lower among those who used LLMs.

The most comprehensive study to date on how AI tools affect cognitive load—including brain activity recordings—was recently published by a group of researchers from MIT (Kosmyna et al., 2025). The study focused on investigating the impact of external technological tools used during the learning process on brain activity. Participants were divided into three groups based on the type of tool they used: those who used large language models (LLMs), those who used traditional search engines, and those who did not use any external tools. Each participant was asked to write essays in four separate sessions using their assigned tools, while their brain activity was monitored throughout the process.

The most striking finding of the study was that while brain activity differed significantly between groups, it remained relatively homogeneous within each group. In other words, brain activity varied according to the use and nature of the external tool employed. The highest levels of brain activity were observed among participants who did not use any external tools, whereas the lowest levels were recorded in those who used LLMs while writing their essays. This finding suggests that the more a given tool demands active participation in cognitive processes, the greater the corresponding brain activity. Accordingly, AI applications—particularly LLMs—elicited the lowest level of cognitive demand in this context.

Another notable finding of the study concerned participants' sense of ownership over their essays and their ability to recall details from them. Participants in the LLM group—which required the least cognitive engagement—reported the lowest levels of ownership toward the essays they had written. Moreover, there was a direct correlation between the sense of ownership and memory recall: as the sense of ownership declined, participants remembered significantly fewer details from the essay they had just written. These findings align closely with those reported by Stadler et al. (2024) regarding cognitive offloading.

When these studies are evaluated collectively, they reveal a clear pattern: when external technological tools used in the learning process demand less active cognitive engagement from individuals, brain activity decreases, cognitive load is reduced, cognitive offloading increases, learning becomes more superficial, critical thinking weakens, and ultimately, the connection to what is produced—i.e., memory—is diminished. This same risk applies in the long term when such tools are used within the education system. Students' overreliance on AI tools pushes them toward passive rather than active participation in fulfilling their academic responsibilities. In this way, the external tool meant to assist them ends up acting as a substitute, but this substitution carries a high cognitive cost. As a result, students not only miss the opportunity to identify and address their own learning gaps—since they mask success as if it were their own—but also experience a gradual decline in their critical thinking and independent problem-solving skills. The weakening of these core skills, in turn, reinforces behavioral dependency on AI tools.

DISCUSSION

Just as AI tools are transforming other sectors, they are rapidly reshaping the education system as well. While these tools offer numerous benefits to students, teachers, and education administrators, the risks they pose—both current and potential—are equally critical. This study has focused on the possible negative impacts of the widespread use of AI tools in education, particularly on students' critical thinking skills.

It is well known that students' overreliance on AI applications often leads them to use these tools as substitutes for their own efforts. Recent studies indicate that this dependency, fueled by excessive trust, distances students from active engagement in the learning process, results in more superficial learning, and weakens memory retention. In fact, this outcome can be seen as a direct consequence of the fundamental nature of AI itself. The core promise of AI tools is to relieve humans of cognitive burdens by taking over certain tasks and producing seemingly reasonable responses quickly. During this process, human cognitive load is reduced. Theoretically, such a reduction should allow individuals to better review the generated content or use the saved time for other productive activities. However, the actual outcome diverges significantly from this expectation. The very feature of AI that reduces cognitive effort encourages students to use it more frequently and with growing reliance. Ultimately, this excessive trust fosters a pattern of behavioral dependency.

As students distance themselves from cognitive processes, they gradually become more passive, increasingly replacing their own decision-making with reliance on AI, and ultimately experience a decline in their ability to solve problems independently. In other words, although AI may appear to liberate students from cognitively demanding and formative learning experiences—those that deepen understanding and ensure lasting learning—it in fact increases the cognitive cost over time, negatively impacting the development of essential mental faculties. This adverse effect continuously reinforces overreliance and leads students to adopt AI-generated biased or inaccurate information—such as hallucinated content—without subjecting it to critical thinking or rational scrutiny. As a result, biases embedded in societal power structures related to religion, culture, race, or gender may be reinforced through AI applications more powerfully and pervasively than ever before—particularly through the education system.

If no measures are taken regarding the use of AI applications, the decline in critical thinking that will likely emerge in the medium and long term may give rise to yet another problem: Who is the true author of essays, assignments, and projects? This issue has already been the subject of debate in the academic community and among journal editorial boards, particularly regarding whether AI tools can be used in the production of scientific articles and, if so, how such contributions should be acknowledged (Özer, 2024b). As articles listing AI as a co-author began to appear, the first point of contention was whether AI could be considered an author at all. At the heart of this debate lay the question of whether AI could assume responsibility. The consensus eventually emerged that AI cannot be an author because it cannot bear responsibility. Although a few scientific journals have prohibited the inclusion of AI-generated texts, graphics, or other content altogether (Thorp, 2023), the majority of editorial boards have instead adopted policies allowing such contributions—on the condition that they are clearly disclosed within the article, as they may enhance the quality of the scientific work. In the context of education, however, the situation becomes even more ethically fraught. Far from disclosing the use of AI, students are more likely to present AI-generated work as their own, raising serious ethical concerns. This kind of distorted moral behavior may lead students to undervalue human effort and contribution. At the same time, a paradox emerges: in a system where education is expected to cultivate critical thinking, the widespread use of AI may instead result in its steady erosion.

In short, AI applications in education possess a dual character: they can both support and suppress critical thinking. If left unchecked, recent studies—including those referenced in this article—clearly indicate that the suppressive dimension is likely to become dominant. Therefore, it is essential to prevent the emergence of the contexts in which this suppressive effect on critical thinking is activated. AI tools must be used in ways that promote cognitive engagement and active interaction with complex content, particularly within educational settings. In other words, generative AI tools should not be evaluated solely based on the convenience they offer, but also on how they are used and what kinds of cognitive processes they stimulate.

In this context, both the intensive individual use of AI tools by students and their broader integration into education systems necessitate a strong and continually reinforced foundation in AI literacy. The primary focus of this literacy should be on how to use these technologies ethically and responsibly. Students, teachers, and education administrators must be made aware of issues related to data security, the potential biases in AI-generated content, and the fact that such content may not always be accurate. They should be encouraged to critically assess and question the outputs produced by AI tools. Of course, while such educational efforts are essential, they are not sufficient on their own. Given that excessive use and overreliance on AI tools can foster dependency behaviors, interventions must also be introduced to prevent this type of misuse in educational settings. These interventions should aim to prevent AI tools from replacing students in the learning process. Instead, strategies must be developed to employ AI as a supportive mechanism that enhances, rather than substitutes, students' active participation in cognitive processes.

Financial disclosure: No financial support was received.

Conflict of interest: There is no conflict of interest.

Author Contributions: Authors have contributed equally to the work

Refereeing Process: Double-blind refereeing.

REFERENCES

Ahmad, S. F., Han, H., Alam, M. M., Rehmat, M. K., Irshad, M., Arraño-Muñoz, M., & Ariza-Montes, A. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications, 10*(311), doi: 10.1057/s41599-023-01787-8

Ambele, R. M., Kaijage, S. F., & Dida, M. A., Trojer, L., & Kyando N. W. (2022). A review of the development trend of personalized learning technologies and its applications. *International Journal of Advances in Scientific Research and Engineering*, 8 (11), 75–91. doi: 10.31695/IJASRE.2022.8.11.9

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Athaluri, S. A., Manthena, S. V., Kesapragada, V. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus*, *15*(4). doi: 10.7759/cureus.37432.

Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. Retrieved from https://digitalcommons.uri.edu/cba_facpubs/548.

Berberette E, Hutchins J, Sadovnik A. Redefining "Hallucination"in LLMs:Towards a psychology-informed framework for mitigating misinformation. arXiv. 2024:2402–01769v1.

Chiu, T. K. F. (2023). The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10),6187-6203. doi: 10.1080/10494820.2023.2253861

Ghosh, S., Venkit, P. N., Gautam, S., Wilson, S., & Caliskan, A. (2024). Do generative AI models output harm while representing non-Western cultures: Evidence from a community-centered approach. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (Vol. 7, pp. 476-489).

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1). doi: 10.1126/sciadv.aao5580

Gangadharbatla, H. (2022). The role of AI attribution knowledge in the evaluation of artwork. *Empirical Studies of the Arts*, 40(2), 125–142. doi: 10.1177/0276237421994697

Gerlich, M. (2025). AI Tools in Society: Impacts on cognitive offloading and the future of critical thinking, *Societies,* 15(1), 6. doi: 10.3390/soc15010006

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. doi: 10.3390/educsci13070692

Heaton, D., Nichele, E., Clos, J., & Fischer, J. E. (2023). "The algorithm will screw you": Blame, social actors and the 2020 A Level results algorithm on Twitter. *PLoS One,* 18(7). doi: 10.1371/journal.pone.0288662

Idowu, J. A. (2024). Debiasing education algorithms. *International Journal of Artificial Intelligence in Education, 34*, 1510–1540. doi: 10.1007/s40593-023-00389-4

Ilikhan, S., Özer, M., Tanberkan, H., & Bozkurt, V. (2024). How to mitigate the risks of deployment of artificial intelligence in medicine? *Turkish Journal of Medical Science*, 54(3), 483-492. doi: 10.55730/1300-0144.5814

Ilikhan, S., Özer, M., Perc, M., Tanberkan, H., & Ayhan, Y. (2025). Complementary use of artificial intelligence in healthcare. *Medical Journal of Western Black Sea*, 9(1), 7-17. doi: .29058/mjwbs.1620035

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38. doi: 10.1145/3571730

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D.,Fischer, F.,…Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*. doi: 10.1016/j.lindif.2023.102274

Khan, R. A., Jawaid, M., Khan, A. R., Sajjad, M. (2023). ChatGPT—Reshaping medical education and clinical management. *Pakistan Journal of Medical Sciences*, 39(2), 605–607. doi: 10.12669/pjms.39.2.7653

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V.,… Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv:2506.08872*. doi: 10.48550/arXiv.2506.08872

Lum, K., & Isaac, W. (2016). To predict and serve?. *Significance*, 13(5), 14–19. doi: 10.1111/j.1740-9713.2016.00960.x

Mohammadkarimi, E., & Omar, J. A. (2025). Does artificial intelligence impede critical thinking? A case study of Iranian university students. *Journal of Applied Learning & Teaching*, 8(2), 1-9. doi:10.37074/jalt.2025.8.2.3

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453.doi: 10.1126/science.aax2342

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Books.

Özer, M. & Suna, H. E. (2022). The relationship between school socioeconomic composition and academic achievement in Turkiye. *Journal of Economy Culture and Society*, 66, 17–27. doi: 10.26650/JECS2022-1116979

Özer, M. (2024a). Potential benefits and risks of artificial intelligence in education. *Bartın University Journal of Faculty of Education*, 13(2), 232-244. doi: 10.14686/1416087

Özer, M. (2024b). Impact of ChatCPT on Scientific Writing. *The Journal of Humanity and Society*, 14(3), 210-217. doi: 10.12658/E0002

Özer, M. (2024c). Is artificial intelligence hallucinating?. *Turkish Journal of Psychiatry*, 35(4), 333-335. doi: 10.5080/u27587

Özer, M. (2024d). Dynamics of the meritocracy trap and artificial intelligence. *International Journal of Management Economics and Business,* 20(3), 845-869. doi: 10.17130/ijmeb.1524229

Özer, M., & Perc, M. (2024). Human complementation must aid automation to mitigate unemployment effects due to AI technologies in the labor market. *Reflektif Journal of Social Sciences*, 5(2), 503-514. doi: 10.47613/reflektif.2024.176

Özer, M., Perc, M., & Suna, H. E. (2024a). Artificial intelligence bias and the amplification of inequalities in the labor market. *Journal of Economy, Culture and Society*, 69, 159-168. doi: 10.26650/ JECS2023-1415085

Özer, M., Perc, M., & Suna, H. E. (2024b). Participatory management can help AI ethics adhere to the social consensus. *Istanbul University Journal of Sociology*, 44(1), 221-238. doi: 10.26650/SJ.2024.44.1.0001

Özer, M. (2025). Can mathematical models be weapons of mass destruction? *Reflektif Journal of Social Sciences,* 6(1), 259-268. doi: 10.47613/reflektif.2025.212

Perc, M., Özer, M., & Hojnik, J. (2019). Social and juristic challenges of artificial intelligence. *Palgrave Communications*, 5(61). doi: 10.1057/s41599-019-0278-x

Qadri, R., Shelby R., Bennett, C. L., & Denton, R. (2023). AI's Regimes of Representation: A Community-centered Study of Text-to-Image Models in South Asia. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency,* 506–517. doi: 10.1145/3593013.3594016

Rudolph, J., Tan, S., & Tan, S. C. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 1-22. doi: 10.37074/jalt.2023.6.1.9

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance. *Psychological Bulletin, 135*(1), 1–22.

Said, E. (1979). *Orientalism*. New York:Vintage Books

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior,* 160. doi: 10.1016/j.chb.2024.108386

Suleyman, M. (2023). *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma*. New York: Crown.

Suna, H. E., Tanberkan, H., Gür, B. S., Perc, M. & Özer, M. (2020). Socioeconomic status and school type as predictors of academic achievement. *Journal of Economy Culture and Society*, 61, 41-64. doi: 10.26650/JECS2020-0034

Suna, H. E. & Özer, M. (2021). The impact of school tracking on secondary vocational education and training in Turkey. *Hacettepe University Journal of Education*, 36(4), 855-870. doi: 10.16986/HUJE.2021068158

Suna, H. E., Özer, M. Sensoy, S., Gür, B. S., Gelbal, S., & Aşkar, P. (2021). Determinants of academic achievement in Turkey. *Journal of Economy Culture and Society*, 64, 143-162. doi: 10.26650/JECS2021-934211

Suna, H. E. & Özer, M. (2022). The relationship of preschool attendance with academic achievement and socioeconomic status in Turkey. *Journal of Measurement and Evaluation in Education and Psychology*, 13(1), 54-68. doi: 10.21031/epod.1060460

Suna, H. E. & Özer, M. (2024). Medium- and long-term outcomes of early childhood education: experiences from Turkish large-scale assessments. *Humanities and Social Sciences Communications*, 11, 853. doi: 10.1057/s41599-024-03241-9

Suna, H. E., & Özer, M. (2025). The human complimentary usage of AI and ML for fair and unbiased educational assessments. *Chinese/English Journal of Educational Measurement and Evaluation*, 6(1), 1-21. doi: 10.59863/YPKL4338

Tanberkan, H., Özer, M., & Gelbal, S. (2024). Impact of artificial intelligence on assessment and evaluation approaches in education. *International Journal of Educational Studies and Policy*, 5(2), 139-152. doi: 10.5281/zenodo.14016103

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313. doi: 10.1126/science.adg7879

Tsai, S. C., Chen, C. H., Shiao, Y. T., Ciou, J. S., & Wu, T. N. (2020). Precision education with statistical learning and deep learning: A case study in Taiwan. *International Journal of Educational Technology in Higher Education*, 17, 12. doi: 10.1186/s41239-020-00186-2

Ulnicane, I., & Aden, A. (2023). Power and politics in framing bias in artificial intelligence policy. *Review of Policy Research*, 40 (5), 665–687. doi: 10.1111/ropr.12567

Villegas-Ch, W., Arias-Navarrete, A., & Palacios-Pacheco, X. (2020). Proposal of an architecture for the integration of a Chatbot with artificial intelligence in a smart campus for the improvement of learning. *Sustainability*, 12(4), 1500. doi: 10.3390/su12041500

Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D.,…Tu, Z. (2023). Document-level machine translation with large language models. *arXiv:2304.02210.* doi: 10.48550/arXiv.2304.02210

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1-27. doi: 10.1186/s41239-019-0171-0

Zhang M, Press O, Merrill W., Liu, A., Smith, N. A. (2023). How language model hallucinations can snowball. *arXiv:2305–13534.* doi: 10.48550/arXiv.2305.13534

Zylinska, J. (2023). Art in the age of artificial intelligence. *Science*, 381(6654), 139-140. doi: 10.1126/science.adh0575