

Neurotechnologies in education: personalization of learning through a recommendation service

Tatyana V. Bukina

*Baltic Center for Neurotechnology and Artificial Intelligence
Immanuel Kant Baltic Federal University
Kaliningrad, Russia
Research Institute of Applied Artificial Intelligence
and Digital Solutions
Plekhanov Russian University of Economics
Moscow, Russia
bukinatatyanav@gmail.com*

Marina V. Khramova

*Baltic Center for Neurotechnology and Artificial Intelligence
Immanuel Kant Baltic Federal University
Kaliningrad, Russia
Research Institute of Applied Artificial Intelligence
and Digital Solutions
Plekhanov Russian University of Economics
Moscow, Russia
mkhramova1@kantiana.ru*

Abstract—This article examines the phenomenon of automation bias as one of the significant cognitive biases accompanying the introduction of artificial intelligence (AI) systems into education. An analysis of the scientific literature has been carried out in order to determine the mechanisms of this distortion, its manifestations and associated risks. The main threats associated with the use of AI in such areas as automated knowledge assessment, personalization of training and administrative management have been identified. The paper suggests ways to mitigate risks aimed at building adequate trust in technology.

Keywords—artificial intelligence, education, automation bias, cognitive bias

I. INTRODUCTION

The integration of artificial intelligence (AI) systems into the modern educational environment is one of the most significant transformations in pedagogical science and practice. AI technologies have the potential to radically change the educational landscape, offering opportunities for personalizing learning, automating routine administrative tasks, and improving the overall effectiveness of the educational process [1, 2]. The potential of AI, from adaptive tutor systems to forecasting tools, seems significant for solving a number of urgent pedagogical tasks [3, 4]. It is assumed that automation can free up to 40% of a teacher's time, which will allow resources to be directed to deeper pedagogical interaction with students [1].

Nevertheless, against the background of technological optimism, concerns about the fundamental ethical, psychological and pedagogical consequences of the introduction of AI systems are increasing [5, 6]. Issues related to the bias of algorithms, the preservation of critical thinking among students, and the transformation of human interaction in the learning process are coming to the fore [7].

At the center of this problem lies the complex interaction in the human-AI system. The effectiveness of AI-based educational technologies is determined not only by the perfection of algorithms, but also by how participants in the educational process perceive and interpret the information they generate. A phenomenon that requires close attention is automation bias, a cognitive bias, one of the manifestations of which is unreasonable confidence in the correctness of decisions provided by automated systems [8]. This distortion

can not only negate the advantages of AI, but also lead to systematic errors when using such systems in education.

II. THEORETICAL ASPECTS OF AUTOMATION BIAS

First of all, let's analyze the mechanisms of automation bias and the factors influencing its manifestation.

Automation bias appears to be a tendency to rely on automated prompts instead of carefully searching and processing information. This phenomenon is not an accidental error, but a systematic cognitive bias, a stable model of thinking [9]. The manifestations of automation bias are classified into two types of errors [8]:

1) *Commission errors* it is an incorrect actions performed by the user as a result of blindly following the erroneous recommendation of the system.

2) *Omission errors*—the lack of necessary actions on the part of the user, since the system did not provide the appropriate signal.

This cognitive bias is based on a number of fundamental psychological mechanisms. In particular, the human brain tends to minimize mental effort, and in conditions of multitasking, automated systems are perceived as a way to reduce cognitive load [10]. In addition, there is a widespread belief that technology is inherently more objective and reliable than humans, which creates an "aura of infallibility" around AI's conclusions [11].

B. Differentiation of related concepts

In order to correctly diagnose the problems that arise during the implementation of AI, it seems important to make a terminological distinction between automation bias and two related phenomena: algorithmic bias and automation complacency.

1) *Algorithmic bias* it is a characteristic of the system itself and represents systematic errors in the algorithm, leading to unfair results for certain groups [12].

2) *Automation complacency*—a state of reduced vigilance and insufficient monitoring of the system as a result of high confidence in its reliability [8].

Below is a comparative table showing the differences between these concepts (Table I).

This work is supported by the Russian Science Foundation (project 23-72-10016).

TABLE I. COMPARATIVE ANALYSIS OF RELATED CONCEPTS

Parameter	Concepts		
	<i>Automation bias</i>	<i>Algorithmic bias</i>	<i>Automation complacency</i>
Source	Cognitive Bias	Technical/statistical defect	Reduced monitoring
The main manifestation	Uncritical acceptance of recommendations	Systematically unfair results	Passive monitoring
An example in education	The teacher accepts an erroneous essay assessment from the AI without double-checking it	The AI system for university admission systematically rejects applications from applicants from certain areas	The student does not notice that the adaptive learning system has stopped offering new materials, as he is used to it always working correctly
Key risk mitigation strategy	User training, interface design, explicable AI	Audit and correction of data/algorithms	Dynamic task allocation

III. RISK ANALYSIS IN KEY AI APPLICATIONS

The theoretical risks associated with automation bias acquire concrete manifestations when they are considered in key areas of educational activity.

A. Automated assessment and feedback

One of the high-risk areas of AI application is the automated assessment of training tasks. A teacher who is subject to automation bias can completely delegate the verification process to the system, as a result of which errors and biases of the algorithm are fixed in the form of an official assessment. Students, in turn, may show uncritical trust in feedback, which leads them to form misconceptions about their academic competencies.

B. Personalized Learning Systems

Adaptive educational platforms offering the construction of individual educational trajectories for each student are another high-risk area. The main risk here is to reduce the student's independence. With full confidence in the "optimal" path proposed by the system, the student risks transforming from an active subject of his own learning, capable of research and reflection, into a passive recipient following the instructions of the algorithm. Equations

C. Long-term impact on cognitive skills

In addition, excessive reliance on AI can lead to a gradual deterioration in the ability to think independently. Two main mechanisms of negative influence are assumed:

1) *The erosion of critical thinking.* When AI provides ready-made answers and generalizations, students lessen the need to practice information analysis skills, evaluate arguments, and form informed conclusions [13].

2) *Cognitive offloading* describes the process of transferring mental effort to external tools. In the long term, this prevents the formation of stable neural connections and deep mental models, which can lead to atrophy of the corresponding mental abilities [14, 15].

IV. RISK MITIGATION STRATEGIES AND RESPONSIBLE USE OF AI

The problem of automation bias is multidimensional and requires an integrated approach to its solution. Mitigation strategies should cover technological, pedagogical, and organizational levels, forming a system in which AI is used as a tool to enhance rather than replace human intelligence.

A. Technological and design solutions

Changes at the design level of the AI systems themselves can play a key role in calibrating user trust and mitigating risks. The main approaches include:

1) *Transparency and explicable AI.* Designing systems capable of providing explanations of their decisions in a form understandable to humans, allows the user to better understand the logic of the algorithm. Nevertheless, it should be borne in mind that the presence of explanations is not a panacea and in some cases may paradoxically strengthen automation bias [9].

2) *Implementation of confidence indicators.* Integration of visual or numerical indicators into the user interface, reflecting the degree of "confidence" of the model in its conclusion, can serve as a signal for the teacher about the need for more careful control [16].

3) *Maintaining user control.* The system design should support the active role of the user in the decision-making process. Instead of giving a single "correct" answer, the system can offer several options, require the user to explicitly confirm critical decisions, or provide tools to correct the results [16].

B. Educational strategies for the formation of AI competence

Technological solutions should be complemented by targeted pedagogical efforts to develop AI literacy among participants in the educational process.

1) *Education and training.* Educational institutions should develop and implement training programs aimed at developing a deep understanding of the principles of AI, its fundamental limitations, the probabilistic nature of conclusions and ethical principles [3].

2) *Formation of a critical assessment.* The key task of a teacher in the age of AI is to teach students how to use these tools not as a source of true knowledge, but as an intellectual partner. The role of the teacher is being transformed from a translator of knowledge into the role of a facilitator who helps students develop critical thinking skills and responsible interaction with information in a digital environment [14].

C. Organizational measures

To create a safe environment for using AI, it is necessary to adhere to the rules and standards at the level of educational institutions.

1) It is necessary to introduce *clear rules* for the use of AI in educational institutions, defining the requirements for responsibility for the operation of systems [7].

2) *The Human-in-the-Loop principle.* For all critical decisions that have serious consequences for students (for example, final assessment, enrollment), the principle of mandatory human control should be established. AI can be

used as a support tool, but the final decision should always be left to the individual [17].

3) *Independent audit*. It is necessary to introduce the practice of regular auditing of used AI systems. Such an audit should evaluate systems not only for technical accuracy, but also for algorithmic bias [18].

It is the integrated application of these strategies that will create a sustainable ecosystem for the use of AI in education.

V. CONCLUSIONS

The conducted analysis of scientific literature allows us to conclude that automation bias is a serious and underestimated risk accompanying the integration of AI systems into the educational environment. This cognitive bias, rooted in the fundamental mechanisms of human thinking, transforms AI from a tool for improving efficiency into a source of systemic errors in the long run.

In our opinion, awareness of these risks should not lead to the abandonment of the use of technology. A paradigm shift towards a human-centered approach is needed [6]. Within the framework of the new paradigm, AI should be considered solely as a tool to enhance and support human intelligence. The implementation of this approach requires comprehensive and coordinated efforts at all levels: from developers, who must create transparent and controlled systems, to educators and administrators, who must implement clear ethical frameworks to ensure that the final decision on critical issues always rests with the individual. Only with this approach is it possible to use the potential of artificial intelligence to build a more effective and meaningful education.

ACKNOWLEDGEMENTS

This work is supported by the Russian Science Foundation (project 23-72-10016).

REFERENCES

- [1] M. Martínez-Comesaña et al., "Impact of artificial intelligence on assessment methods in primary and secondary education: systematic literature review," *Revista de Psicodidáctica* (English ed.), vol. 28, no. 2, pp. 93–103, 2023, doi: 10.1016/j.psicoe.2023.06.002.
- [2] J. Liang, J. M. Stephens, and G. T. L. Brown, "A systematic review of the early impact of artificial intelligence on higher education curriculum, instruction, and assessment," *Frontiers in Education*, vol. 10, Art. no. 1522841, 2025, doi: 10.3389/educ.2025.1522841.
- [3] T. V. Bukina, "Artificial intelligence in education: current state and development prospects," *Society: Sociology, Psychology, Pedagogy*, [Т. В. Букина, "Искусственный интеллект в образовании: современное состояние и перспективы развития," *Общество: социология, психология, педагогика*], no. 1 (129), pp. 76–83, 2025, doi: 10.24158/spp.2025.1.9.
- [4] V. V. Grubov et al., "Open-loop neuroadaptive system for enhancing student's cognitive abilities in learning," *IEEE Access*, vol. 12, pp. 49034–49049, 2024, doi: 10.1109/ACCESS.2024.3383847.
- [5] A. Nguyen, H. N. Ngo, Y. Hong, and Y. H. V. Ho, "Ethical principles for artificial intelligence in education," *Education and Information Technologies*, vol. 28, pp. 4221–4241, 2023, doi: 10.1007/s10639-022-11316-w.
- [6] United Nations Educational, Scientific and Cultural Organization, *Recommendation on the ethics of artificial intelligence*, 2022. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- [7] A. M. Vieriu and G. Petrea, "The impact of artificial intelligence (AI) on students' academic development," *Education Sciences*, vol. 15, no. 3, Art. no. 343, 2025, doi: 10.3389/educsci15030343.
- [8] K. Goddard, A. Roudsari, and J. C. Wyatt, "Automation bias: a systematic review of frequency, effect mediators, and mitigators," *Journal of the American Medical Informatics Association*, vol. 19, no. 1, pp. 121–127, 2012, doi: 10.1136/amiajnl-2011-000089.
- [9] M. Vered, T. Livni, P. D. L. Howe, T. Miller, and L. Sonenberg, "The effects of explanations on automation bias," *Artificial Intelligence*, vol. 322, Art. no. 103952, 2023, doi: 10.1016/j.artint.2023.103952.
- [10] D. Lyell and E. Coiera, "Automation bias and verification complexity: a systematic review," *Journal of the American Medical Informatics Association*, vol. 24, no. 2, pp. 423–431, 2017, doi: 10.1093/jamia/ocw105.
- [11] A. Graefe, M. Haim, B. Haarmann, and H.-B. Brosius, "Readers' perception of computer-generated news: credibility, expertise, and readability," *Journalism*, vol. 19, no. 5, pp. 595–610, 2016, doi: 10.1177/1464884916641269.
- [12] A. Jonker and J. Rogers, "What is algorithmic bias?," IBM, Sep. 20, 2024. [Online]. Available: <https://www.ibm.com/think/topics/algorithmic-bias>
- [13] Elsevier, "Shaping an AI-driven future," in *Insights 2024: Attitudes toward AI*, Chapter 3, 2024. [Online]. Available: <https://www.elsevier.com/insights/attitudes-toward-ai/shaping-an-ai-driven-future>
- [14] A. R. Chow, "ChatGPT may be eroding critical thinking skills, according to a new MIT study," *TIME Magazine*, 2025. [Online]. Available: <https://time.com/7295195/ai-chatgpt-google-learning-school/>
- [15] N. Kosmyna et al., "Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task," 2025, arXiv:2506.08872. [Online]. Available: <https://arxiv.org/abs/2506.08872>
- [16] J. Zerilli, U. Bhatt, and A. Weller, "How transparency modulates trust in artificial intelligence," *Patterns*, vol. 3, no. 4, Art. no. 100455, 2022, doi: 10.1016/j.patter.2022.100455.
- [17] D. Braun, P. Rogetzer, E. Stoica, and H. Kurzahls, "Students' perspective on AI-supported assessment of open-ended questions in higher education," in *Proceedings of the 15th International Conference on Computer Supported Education*, vol. 2, pp. 73–79, 2023, doi: 10.5220/0011648900003470.
- [18] Schiller International University, "Risks of AI algorithmic bias in higher education," 2025. [Online]. Available: <https://www.schiller.edu/blog/risks-of-ai-algorithmic-bias-in-higher-education/>