



Copyright © 2024 by Cherkas Global University
All rights reserved.
Published in the USA

European Journal of Contemporary Education

E-ISSN 2305-6746

2024. 13(4): 638-648

DOI: 10.13187/ejced.2024.4.638

<https://ejce.cherkasgu.press>

IMPORTANT NOTICE! Any copying, reproduction, distribution, republication (in whole or in part), or otherwise commercial use of this work in violation of the author's rights will be prosecuted in accordance with international law. The use of hyperlinks to the work will not be considered copyright infringement.



**European Journal of
Contemporary Education**



ELECTRONIC JOURNAL

How Students Develop Creative Skills and Critical Thinking with ChatGPT: Experimental Results

Ilya Bykov ^{a, *}, Irina Gladchenko ^b

^a St Petersburg State University, Russian Federation

^b Moscow State University, Russian Federation

Abstract

Mass access to AI-technologies like ChatGPT or other AI-assisted tools has significantly expanded the student's ability to search for and process vast amounts of information during the education process. Student use AI-generated texts to develop projects in the human sciences. It seems that application of synthetic content into student work is inevitable. This paper aims to study development of students' critical thinking and creative skills with application of ChatGPT. Haphazard delegation of the initiative to AI-based chatbots directly affects the development of these skills. To evaluate these effects an experimental approach has been applied. Sixteen students have participated in the experiment. There were two groups. Each group prepared weekly materials on a predetermined topic. The experimental group could use only AI-based chatbots. The control group could use any other resources, except chatbots. Students had written an initial essay and final abstracts for the conference. After the experiment, students provided feedback about working with AI. The results showed that critical thinking parameters and creative skills have been comparatively underdeveloped in the experimental group.

Keywords: AI, media literacy, critical thinking, creative skill, higher education, ChatGPT, international relations, political science, political communication.

1. Introduction

The number of sophisticated technologies that mankind uses to improve its standard of living grows rapidly. Explosive technological growth provokes public discussion about the place of technologies in our future. How could new instruments change our daily lives and social structure? What is the place of a human being in new social paradigm? New machines, robotic technology and AI-based virtual bots are firmly in place now. What is the fundamental difference between human and machine activity? Scientists also raise concerns that AI could soon take a leading role in the

* Corresponding author

E-mail addresses: i.bykov@spbu.ru (I. Bykov), irinaglad94@mail.ru (I. Gladchenko)

social structure. They call for an active debate about how a human differs from AI and how to integrate it into everyday life (Van Quaquebeke, Gerpott, 2023).

All of these issues are related to the educational process (Bykov, Medvedeva, 2024). It is especially important to know what skills should be in demand for development of the next generations and who will integrate innovations into their daily lives. AI-based programmes are able to perform algorithmic tasks much faster than humans. AI-technologies are convenient for a number of manufacturing tasks. This is why AI-based technologies are already finding applications in a variety of human endeavours. This situation inevitably leads to the use of human resources' reconsideration in different industries. Organisation of warehouses' distribution systems and development of unmanned transport could be a good illustration. If robotic technologies can handle algorithmic activities better than humans typically do, then what should be the direction of human endeavour? For example, Ray notes that ChatGPT's IQ, based on various tests, ranges from 147 to 150 points, which is a very high figure. A language model can pass a number of tests and even pass some exams (Ray, 2023: 134).

Corresponding examples can also be found in social and humanitarian spheres. In 2023, there were attempts to replace news anchors with deepfakes (Deepfake..., 2023). The potential of using AI is already being used quite boldly. The lack of clear legislative framework and control mechanisms also contribute to the use of AI-technologies in virtual communication. For example, characters generated by artificial intelligence could easily enter social virtual space. A human could distinguish a fake from a real person now. But it is quite possible that in the future a real person on the video will be practically indistinguishable from a virtual one (Bykov, Medvedeva, 2024). A flow of events and facts around modern man is currently so huge that it is impossible to control it. That is why development of critical thinking, an ability to verify information and to assess critically surrounding events are becoming a necessary part of social life (Abrami et al., 2015; Darwin et al., 2023; Essien et al., 2024; Hejres, 2022; Nassar, 2019; Spector, Ma, 2019). Educational projects which chatbots partially replacing teachers are under development. Some studies also indicate effectiveness of utilising technologies in educational activities (Bykov et al., 2019; Faqih, 2023; Gui, Lee, 2023; Walter, 2024; Yi, 2024).

The penetration of AI technologies into social processes is already deep enough to start testing the limits to what extent it is possible to use AI without human control. Spheres with active interpersonal interaction are particularly vulnerable in this regard. One of such areas is education. Educational activities imply not only training of technical skills to wield various researches or using production tools but also learning how to critically and creatively perceive reality and making sense of it. Development of critical perception is an integral part of an educational process. Perhaps, an ability to think critically will be the key aspect that will distinguish human activity from that of artificial intelligence. On the other hand, it is equally important to develop students' skills for creative production and original thinking.

Modern studies of critical thinking are based on Bloom's taxonomy. Bloom offered several categories of which critical thinking is composed. Among these categories are "knowledge, comprehension, application, analysis, synthesis, and evaluation" (Bloom et al., 1956). These categories include subcategories, clarifying basic ones. Also, subcategories are ranged from more concrete to more abstract. Thus, "knowledge" includes knowledge of facts, terminology, trends, classifications, methodology, theories, structures, etc. Nevertheless, this subcategory implies only the existence of knowledge as a fact. The "comprehension" section includes abilities interpretation and extrapolation, among others. "Analysis" implies an ability to analyse relations, elements, and principles. Category "synthesis" is the most interesting in terms of current research and includes an ability to produce new knowledge and make a conclusion (Bloom et al., 1956). In an overview on Bloom's taxonomy, Krathwohl relabels "synthesis" into "creativity" (Krathwohl, 2002). Also, the comprehension of critical thinking could have a structure of five stages: "absolute beginner (novice), advanced beginner (beginner), competent performer (competent), proficient performer (proficient), and intuitive expert (expert)" (Spector, Ma, 2019). A meta-analysis taken in 2015 shows that critical thinking should be viewed as a set of skills rather than a separate monolithic skill (Abrami et al., 2015). Recent studies also show that scientists have begun to consider not only understanding and strict adherence to methodology but also questioning skills and healthy scepticism as part of their analysis of critical thinking (Darwin et al., 2023).

The proposed classification allows quite clearly to define skills that a student should have after graduation. Nevertheless, we modified the classification and adapted it to our specific task.

According to Bloom's taxonomy, the main part is devoted to direct knowledge work: accumulation, memorization, searching for regularities, producing results. Almost all of these skills could be already detected in materials produced by AI-powered programmes. In this study, focus also was on creativity. Some aspects of critical thinking are already implicated into AI-powered programs because they are well-structured and rational enough for a computer program. However, creativity implies an irrational approach that is still a mystery to scientists. Its integration into AI-powered programs has not yet happened. Therefore, we were also interested in comparing the aspects of critical thinking and creativity development in students who used AI programs and standard learning resources. To determine the level of creativity, we appealed to Guilford's classification (Guilford, 1973).

This paper aims to study development of students' critical thinking and creative skills with application of ChatGPT. We intend to explore whether a student can produce new knowledge only by accumulating what already exists. In this study we did not analyse students' progress on the basis of all five stages of critical thinking. The aim was to test whether there would be any changes at all. That is why the purpose of the experiment was to observe the changes of undergraduate students' critical thinking and creative skills when they use an AI-based chatbot as their only information source. We decided to conduct an experiment to understand how artificial intelligence technologies can be utilised in educational activities without guidance from the university's instructor.

2. Materials and methods

The method applied in this study belongs to the group of qualitative research methods. Qualitative methods typically involve smaller sample sizes and focus on depth rather than breadth. Researchers collect data through various techniques, including interviews, focus groups, observations, experiments, and content analysis. These methods allow for a more nuanced understanding of participants' perspectives, capturing the complexity of their experiences.

The experiment was conducted for ten weeks from February to May of 2024. The undergraduate students of the Moscow State University majoring in international relations participated in the experiment. In the experiment took part students of the second and third years. All students of the second and third year participated in the experiment. Since the experiment took quite a long time and required regularity in checking homework completion and participants' voluntary consent to publish the results, only 16 students gave their consent to publish the results. Results obtained on a larger volume correlate with results presented in the paper. Students of the course are divided into equal groups. Groups are formed in the first year and do not change during the whole period of the bachelor's program. Group formation is initially influenced by language level. The number of students in the group did not change during the selected semester. For the experiment, within their groups, the students were divided into two parts. Students had equal language level within each group. The language level of the students whose results were published was equal. Language level had no significant effect on the experiment results.

The control group did their homeworks using all available "traditional" sources, i.e. printed books, online resources, classmates' or friends' knowledge, social media channels, search engines, electronic dictionaries, paper dictionaries, etc. The experimental group could use only AI-powered chatbots to prepare their tasks. Students could use any publicly available chatbots. It was also allowed to use multiple AI-based resources at the same time. Students independently chose a way to accomplish a task during semester. Under conditions of the experiment students were required to submit their homework to a lecturer every week without skipping. An entrance and a final essay were required to participate in the experiment. If a student missed at least one task, he/she was excluded from the experiment. Once a method of completing an assignment ("AI" or "traditional") was selected during the semester, that method could not be changed. The course was taught in English. The program focuses on communicative aspect of international relations. The program includes modules on "propaganda", "information warfare", "climate change", "culture", "soft power", "interest groups and lobbying".

The research was conducted in three stages: 1) writing the entrance essay, 2) submitting tasks on given topics, 3) writing the final text in conference abstract format. The task of the entrance essay was to assess the students' baseline. For a moment of writing entrance essay was important that students had not been practicing skills being tested for a significant period. The time period of one winter month implies a break in learning activities. From our point of view, that is sufficient to resume the study process with available so-called residual knowledge. Among the requirements for

essay's structure were the next: 1) to formulate a problem; 2) to provide an introduction to the problem; 3) to express one's thoughts about the problem (to supplement the background knowledge with one's own thoughts). The first part allowed us evaluate the clarity in problem statement. The second part revealed student's general background knowledge. At last, the third part helped to estimate student's originality of thinking and inference skills expressed through ability to supplement background knowledge with one's own thoughts. There was no limit to the number of words. Nevertheless, it was necessary to keep the structure, which consisted of introduction, main part, conclusion. Essay topics were formulated on the basis of thematic sections to be mastered in the second and third year program. Students had two academic hours for writing the essay. It was not allowed to use auxiliary materials, as well as to communicate with classmates.

The second stage was the longest one. During the second stage, students were asked to complete assignments within the framework of the course program. The assignments had to be written and submitted weekly. Students could submit handwritten version, printed version or send to a lecturer an electronic version. There were two formats for assignments: an essay and a report. The report implied the preparation of a general overview of the topic. There were no limits on the number of lines, paragraphs, graphs, pictures, tables, and other textual and visual material.

The final assignment also had two academic hours to complete. It was not allowed to use auxiliary materials as well. However, it was possible to consult with classmates on any issue. The length of the text was limited to 500 words. The task was to write final texts on the given topics in the format of theses for a conference. One main topic with subtopics was proposed for each group. In each group, the lecturer chose a topic that suited the group. Students could choose subtopics from the proposed ones, or they could formulate their own. The main requirement was that the subtopic should respond to the main topic. Students had had to cover four structure elements: actuality of the issue, a problem, a research question, and propose a method of solution.

Description of question actuality allows to assess general background knowledge about chosen research subject. Students were introduced to text structure requirements. Lecturers outlined for students that the first section should contain coverage of current situation in selected political sphere. In this part we offered students to write about practical side of a chosen political issue and the possibility of applying research results on the question in practice. We did not ask students to describe current research field or provide any scientific literature review. Final assignment supposed to reveal student's erudition degree, understanding of modern international political processes regularities, an ability to identify points of potential growth in the field of world politics. For problem section, lecturers indicated that it was necessary to identify any tension or contradiction that exists between political actors in the topic under discussion. It could also be a problem in the practical field of international relations, if a solution and work on which is important for the political international sphere. Here we wanted to estimate student's ability to formulate a problem clearly. While formulating the research question, students had to select one aspect of the problem. The aspect should have been formulated as a question. We also expected students to understand what problem the question would address, to hypothesise what the final outcome would be, and what type of result should be obtained.

The part of determining problem solution method was the next one. Here we asked students to consider a possible way of finding an answer to the research question stated in previous section. The course programs did not include a focused study of methodology. Nevertheless, some homeworks contained references to sociological research methods like surveys and quantitative data analysis. We introduced this section to assess students' originality of thinking, resourcefulness, and inventiveness. Our goal here was to leave room for imagination, assumptions, and searching for possible problem solutions.

Thus, the experiment was designed to note changes in four parameters: erudition, problem statement, original thinking, inference skills. The first parameter is "erudition". In terms of Bloom's taxonomy (Bloom et al., 1956), it corresponds to the "knowledge" category. The second parameter is the ability to define and articulate the research question and/or a problem. This parameter is consistent with "comprehension". The third parameter is originality of thinking. This parameter determines the ability to highlight non-obvious and not often mentioned and discussed environmental interrelationships and to approach the issue unconventionally. "Original thinking" represents part of the "analysis" category. The fourth parameter refers to student's ability to make an inference based on information presented in his/her text. Also it refers to an ability to add something new to already existing knowledge, i.e. to assume existence of new trends, patterns,

facts, phenomena. Finally, it reveals capacity to propose further directions for research. This is an ability to analyse and evaluate information as an independent and objective researcher (corresponds with an ability to draw conclusions). The last parameter corresponds to “evaluation” and “synthesis”.

We measured required parameters by entrance and final essays. The evaluation was based on a point system. A score of zero (0) indicates that the measured parameter was not detected at all. For “erudition” parameter we gave zero (0) points for demonstration of complete absence of knowledge on the topic. In section “problem statement” zero (0) points appeared in cases of 1) unclear problem statement, 2) too broad/general problem statement. In “original thinking” zero (0) points went to a student for poor or absent ability to summarize information and evaluate it. In “inference skills” column zero (0) points appeared when a student could not demonstrate skills in relevant data compilation and reaching final verdict on the topic.

If it was possible to detect the parameter but it was described only in general or was a recitation of common knowledge, half a point (0,5) was awarded. One (1) point indicates that the parameter is easily defined in the text. It is represented by description or/and considerations, or formulated as a concrete phrase or a sentence. If we identified the parameter as one at the basic level and then it was improved, another half-point (0,5) or one (1) point were added to the initial score. If the parameter has not been changed, the same score is given as at the initial stage. Thus, if a student initially showed a high score and his/her performance did not change, the “start” and “final” columns would have had the same score.

Although the course was conducted in English and also included English grammar, vocabulary development, and speaking skills, changes in language proficiency were not taken into account. This was done for two reasons. The first was that this is a different and more algorithmic type of work. The way the results were measured and the way the study was organised were not suitable for studying language proficiency improvement. The work did not include practicing the exercises and lexical part. In case we wanted to evaluate this aspect in addition to critical thinking and creativity the final task would also require modifications and inclusion of a test part. The second reason, however, was mandatory implementation of the basic program for second and third year students. The main condition of the experiment was absence of a feedback and supervisors' non-interference with homework assignments. It was not so easy to organise, as soon as weekly classes implied mandatory discussion of course topics to some extent. We took that into account in designing the experiment and in planning classes. During offline classes lecturers had been paying attention on learning English, while semantic part of the course was done by the students at home.

3. Discussion

The influence of artificial intelligence on the development of critical thinking and the possibility of AI becoming better than humans in critical thinking have long been of interest to scientists (Fedorov, Levitskaya, 2017). Scientists argue that artificial intelligence cannot surpass and therefore replace human intelligence for several reasons. Firstly, there are spheres where the presence of a human cannot be replaced. As an example, the author cites the field of sales. The programme is unable to persuade a person to buy something because the strict algorithm of questions does not meet the liveliness and instant changes in the human mind and emotional state. Difficulties also arise when it is necessary for the AI to update the current knowledge base (Nassar, 2019). The observation that AI needs to update the database in order to adequately respond to newly arriving challenges is noteworthy. From this point of view, we look at AI as a tool that is no different from an ordinary archive or a standard database. Such a viewpoint is quite relevant. Now, with the rapid increase in the number of AI-based tools, society is divided on how to perceive new technologies. Should technologies that utilize AI-based software be considered as part of social relationships? If so, then we automatically assume that AI is a superior force to humans, if only because of its speed of response. If we perceive AI as a tool, the question of its subjectivity cannot arise. However, it is quite possible that a discussion on this topic will lead to a discussion of a fundamentally new paradigm of social relations.

Some studies show that the implementation of AI in the learning process can be beneficial. For example, implementing AI as a tool for self-directed learning and using AI can improve students' speaking skills as well as relieve anxiety before a presentation (Hapsari, Wu, 2022). An experiment conducted in 2023 for students enrolled in a chemistry course showed that students

who used ChatGPT more often than others showed less imbedded abilities to understand theoretical concepts and to draw logical conclusions. The researchers conclude that over-reliance on ChatGPT may lead to passive consumption of information instead of developing critical thinking (Gou, Lee, 2023). Nevertheless, the use of AI-based tools can help students with different educational, cultural, and personal backgrounds adapt to the learning process. Nevertheless, the limits of AI technology must be considered (Walter, 2024). The problem of individual differences in students is also pointed out by Spector (Spector, Ma, 2019). The researcher emphasizes that AI can help to customize the educational environment and create a more comfortable educational environment.

Discussions about individualization of education are noteworthy. At the moment, there is not enough data to understand how exactly the educational process of the future will be organized. At the moment, the use of AI in the educational process is rather chaotic, unstructured, and unregulated. This can be partly explained by the overall changes that are taking place in the technological and societal environment too fast for society to adapt to these changes. Also, research in this area at the moment may take a relatively short period of time to see a clear enough trend. The point is that the implementation of a new tool may be based on old ideas of what the educational process should look like. A similar example could be seen during the 2020 pandemic. This is the year when most organizations in the world, including educational organizations, shifted to remote working and learning. And, despite the availability of the most modern means of communication and virtual interaction, as well as the existence of meta-universes that provide virtual infrastructure for the most complex interaction, there was no educational and communicative breakthrough.

Interestingly, some authors point out that AI technologies can help with routine automated tasks, while, for example, educational managers and teachers can focus on more creative tasks and working with students (Fullan, 2023). In general, it can be noted that most scientists agree that AI technologies can significantly improve the automatic part of human activity. But until now, a number of professions have been tied to conveyorized, algorithmic activities. There is a growing debate about what the advantage of humans over AI is. This is why a number of scientists have come to the conclusion that it is so necessary to actualize and explore ideas about what critical and creative thinking, leadership skills, and an individual's view of the world are. Researchers also caution against handing over the management initiative and instructional leading to AI (Hejres, 2022). Some studies emphasize the need for intelligent implementation of AI in the classroom as a launching pad for students to find ideas that can be developed and questioned (Yi, 2024).

As for the experiment method as a form of conducting research in society, it is also actively used by researchers. As a rule, such studies are complex and include several different methodologies. It is effective to use both qualitative and quantitative methods in educational experimentation. Research methods include surveys as a quantitative-qualitative method, students' independent work in selected courses, and open-ended questions and interview methods. It is indicated that it is effective to use open-ended questions in surveys. This provokes subjects to share thoughts and insights (Essien et al., 2024). Among the methods noted is the usefulness of interviewing students, as it helps to contextualize the results of the survey through personal experience (Faqih, 2024).

The study has limitations. First, it is worth noting the difficulties in forming the sample. We didn't have the opportunity to sample students in a probabilistic way. Participation in the experiment was voluntary for both students and lecturers. Therefore, we can only publish some of students' results who gave consent to process and publish the outcomes. Also, we have tested our hypothesis only among students of one faculty and one field of study, i.e. "international relations". These aspects limit a possibility to generalize the results of the study. Limitation of the study also lies in the fact that the results section does not use statistical analysis methods. Although they are seen as redundant in our current study, development of the topic will allow inclusion of complex statistical methods in a research design to improve the accuracy of obtained results.

It is also worth noting the approach to working directly with AI in ongoing research. As a rule, it is one resource and one specific model, for example, ChatGPT4. Or it may be a particular mode, e.g., a conversation mode with the AI. In this study, the students did not have any predetermined AI model that all students had to use. The study was not intended to focus on a particular resource but on the mode of information as such. For this reason, students were asked to choose the resource they would use and were allowed to use more than one resource. The study used a point system to identify progress or regression in students. No separate interview was conducted. But the

students were asked to write a post-course paper in which they could give a free-form summary of their thoughts on the usability of artificial intelligence.

4. Results

Sixteen people participated in the study. Eight people used AI; eight people used any other resources for homework. All data were anonymized for the study. The table shows the results of the comparison of initial and final parameters. Significant changes in the demonstrated results are highlighted in dark green. Insignificant changes (indicators decreased or increased within half a point) are highlighted in pale green and orange. Significant deterioration of indicators is highlighted in red. Of the 16, only two students kept scores unchanged. Both students used traditional resources to complete their homework. As to the remaining 14 students, the parameters have been changed.

Table 1. Comparison of students' results on the characteristics of critical thinking (N = 16)

Students	Resource type	Erudition		Problem statement		Original thinking		Inference skills	
		Start	Final	Start	Final	Start	Final	Start	Final
Student 1	AI	0,5	0	1	1	0,5	0,5	0,5	0,5
Student 2	AI	0,5	1,5	1	0	1	1	0,5	1
Student 3	AI	0,5	1	0,5	1	0	0,5	0,5	0,5
Student 4	AI	0,5	0,5	0	0	1	0	1	0
Student 5	AI	0,5	0,5	1	1	0	0	0	0,5
Student 6	AI	1	1	1	0,5	1	0	0	0
Student 7	AI	1	1	1	1	0	0,5	0	0
Student 8	AI	0,5	0	0,5	1	0	0	0,5	1
Student 9	other sources	0,5	0,5	0	1	0	1	0	1
Student 10	other sources	1	1	1	0,5	0,5	1,5	0	1
Student 11	other sources	0,5	1,5	1	1	1	1	0,5	0
Student 12	other sources	1	1	1	1	0,5	0,5	0,5	1
Student 13	other sources	1	1	1	1	1	1	1	1
Student 14	other sources	1	1	1	1	1	1	1	1
Student 15	other sources	1	0	1	1	0,5	0,5	0	1
Student 16	other sources	0,5	0	0,5	1	0,5	0,5	0	0

Source: compiled by authors

Thus, the first parameter, 'erudition', showed virtually no change in both groups. Significant improvements by one person in each group were compensated by minor improvements and deteriorations, as well as one significant deterioration in the group that worked with traditional sources. The problem statement parameter showed some improvement in the group that used conventional resources and a slight deterioration in the group that used AI. Significant improvement occurred in the group using traditional resources on the "original thinking" parameter, as opposed to the group using AI.

It is worth noting that the most significant improvement occurred in the 'inference skills' parameter in the group that used conventional sources. Here the results are practically not levelled by a slight deterioration. While in the second group this parameter has not changed practically, there are three insignificant improvements within the possible error and one significant deterioration.

It should be noted that the assignments were completed within the framework of the course. That is, the goals set in the course – improvement of the parameter's "erudition", "problem statement", "original thinking", and "inference skills" – were achieved to some extent. However, homework was the main part of students' work. The working programme allocates the largest number of hours for independent work. The results of the study showed that using only chatbots based on artificial intelligence as a source without feedback from the teacher does not improve these indicators. On the contrary, the use of any other alternative sources, even without feedback from the instructor, allows to achieve a significant improvement of these parameters in students.

In addition to this analysis we assessed also flexibility, fluency, elaboration, tolerance of ambiguity, breadth of interest, sensitivity, independence, and reflection. These parameters were identified by J. Guilford as characteristics of creative adults (Guilford, 1973). Guilford identified 15 characteristics of creativity. We nevertheless included in the analysis only 8 parameters that could be identified in the written text.

Table 2. Comparison of students' results on the characteristics of creativity (N = 16)

Students	Resource type	Flexibility		Fluency		Elaboration		Tolerance of ambiguity		Breadth of interest		Sensitivity		Independence		Reflection	
		Start	Final	Start	Final	Start	Final	Start	Final	Start	Final	Start	Final	Start	Final	Start	Final
Student 1	AI	0	0	0,5	1	0,5	0,5	0	0	0	0	0,5	0,5	0,5	0,5	0,5	0,5
Student 2	AI	1	0,5	0,5	1	0	1	0	0	0,5	0,5	0,5	1	1	1	0,5	1
Student 3	AI	0,5	0,5	0	0,5	0	0	0	0	0	0	0,5	0,5	0,5	0,5	0,5	0,5
Student 4	AI	1	0,5	1	1	0,5	1	1	1	1	1	1	1	1	1	1	1
Student 5	AI	0	0	0,5	0	0	0	0	0,5	0	0	0,5	0,5	0,5	0	0	0,5
Student 6	AI	0,5	1	0	1	0,5	1	0	0,5	0,5	1	0,5	1	0,5	1	0,5	1
Student 7	AI	0	0,5	0,5	1	0,5	1	0,5	1	0	0,5	0,5	1	0,5	0,5	0,5	1
Student 8	AI	0	0	0,5	0,5	0,5	0,5	0	0	0	0	0,5	0,5	0	0	0,5	0,5
Student 9	other sources	0	0,5	0,5	0,5	0,5	1	0	0	0,5	1	0,5	1	0	0,5	0,5	0,5
Student 10	other sources	0,5	0,5	0	0,5	0	0,5	0	0	0,5	1	0,5	0,5	0,5	0,5	0,5	0,5
Student 11	other sources	0,5	0,5	0	1	0	0,5	0	0	0,5	0,5	0,5	1	0,5	1	0,5	0,5
Student 12	other sources	0,5	0,5	1	1	1	1	0	0	0,5	1	1	1	1	1	1	0
Student 13	other sources	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Student 14	other sources	1	1	1	0,5	1	0,5	1	0	1	0,5	1	1	1	1	1	1
Student 15	other sources	0	0	0,5	0,5	0,5	0,5	0	0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
Student 16	other sources	0	0	0,5	0,5	0,5	0,5	0	0	0,5	0	0	0	0	0	0,5	0,5

Source: compiled by authors

The results of creativity parameters analysis show average progress for the majority of students. At the same time, three students showed significant progress in fluency and elaboration. Two of them used AI. Two students who used alternative sources showed regression in the parameters of tolerance of ambiguity and reflection. In general, more progress can be seen in the students who used AI.

In the course of the study, students were told that assignments could be done both in written and printed form. It was not necessary to print out the work; it could be sent electronically. However, only a few papers were handwritten. All others were submitted to the teacher in printed form or sent by mail. It should also be noted that the volume of papers was approximately the same for all students and varied from three to five pages. At the same time, all papers were well structured.

5. Conclusion

There were three important aspects to this study. Firstly, the experiment excluded feedback from the supervisor, which was the teacher, because it was necessary to find out whether it was possible for a student to develop critical thinking skills only through information gathering through available sources, guided only by the framework of topics, without feedback from the supervisor. Secondly, the experiment did not involve practicing standard actions with a given algorithm, such as practicing the use of grammatical constructions or new words when learning a language. The point of the tasks was to collect all the information that could be found on the topic – in the same way the task was set for the students – and to check whether this action without feedback was enough for the student to produce new knowledge only on the basis of the received information. Third, in the behavioural experiment, AI-based chatbots were used as next-generation search engines. Many experiments in the field focus on the formulation of problems, followed by the correction of questions to the chatbots, taking into account recommendations from the AI itself or from teachers. In this case, explanatory feedback was excluded. The work on correcting the prompts, checking the literacy of the received material, and correcting and directing their own knowledge depended only on the student.

The analysis of such parameters as erudition, problem statement, original thinking, inference skills refers to Bloom's taxonomy helping to assess the development of critical thinking skills. In this aspect, the experiment results' show that the use of AI reduces students' critical thinking abilities. In particular, problem statement ability, original thinking and inference skills regress. While the students who used alternative sources significantly increased these same indicators. At the same time, the erudition index is higher in students who used AI.

The assessment of characteristics of creative adults shows an interesting picture. Students who used AI showed progress in several parameters: fluency, elaboration, tolerance of ambiguity, breadth of interest, sensitivity, reflection. According to Guilford's explanations, these parameters indicate students' erudition. Fluency and elaboration refer to the ability to find different possible solutions to a problem and to work out the details of these solutions. Breadth of interest and tolerance of ambiguity refer to interest in abstract ideas and the ability to combine different details of these ideas without contradiction. Sensitivity helps to detect problematic aspects. Whereas

reflection is the skill of seeing the full picture of an issue under study. It is likely that using AI tools to work through a topic helps to broaden the student's horizon and draw attention to aspects that for some reason may not be known or may not be the focus of students' attention. Arguably, AI offers a set of viewpoints that has a lot of variation. This raises the possibility of seeing aspects that were somehow out of focus. This aspect can be compared to what is called forethought capability.

A comparison of the results for the critical and creative thinking skills assessment showed that student 6 showed stable growth, student 13 showed stable results, with student 14 showing significant regression of the result. This regression can be explained by the nature of the final assignment. Unlike the starting essay, the final assignment was more formalized and could provide less space for creativity. However, student 13 even demonstrated high skills in the sense of humor parameter. This item was not included in the analysis because the academic texts did not suggest its presence. Nevertheless, student 13 was creative in completing the assignment. He produced a starter essay in the form of a parody of a New York Times-style newspaper article, following the newspaper layout (arrangement of text and hand-drawn graphic elements) and journalistic style of the newspaper.

After finalizing the experiment, we conducted a free-form survey about semester's work. It is worth noting that when the experimental format was proposed, not all students were enthusiastic about the possibility of submitting tasks with the help of AI. Many were wary of the assignment. After a couple or three weeks of work, several people asked if they could change groups and work with regular sources. However, it was not possible due the experiment rules.

Students were asked to write short essays to express their thoughts about the semester's activity. The students made some interesting points:

- The information that AI-powered chatbots provide is a very general information. It also often requires additional human verification and fact-checking. Also, the information was often incomplete, and some students indicated that it would be good to supplement it with information from other sources.

- Some topics could not be developed through AI using the original wording. In many chats, there is a technical restriction on keywords like “information warfare”, “propaganda”, “war”, “weapon”, and “modern military conflicts”. Nevertheless, the course touched on such topics, and, as students noted, this is a major obstacle for international relations professionals.

- The available open-source free chatbots are not good enough to generate pictures and do not produce infographics. Therefore, almost all reports from the group working with AI are made without pictures and illustrations. A few illustrations in the papers were purely decorative in nature.

- Despite the unreliable data provided, chatbots always provide structured information. It usually fits within 100 words. So, some students entered the prompt ‘keep writing the previous answer’ and got an extra few paragraphs of text. Also, some students used the generated text as a basis and completed some parts of the reports on their own. Also, students used several prompts and assembled a single text from disparate parts.

- Most of the students used many chatbots instead of one.

- It was noted that one of the functions of the AI, which it copes with very well, was very convenient for work: shortening a large text by selecting the main thoughts from it and reducing them to one or two paragraphs. This helped to save time and digest the main thoughts of the text already highlighted by chat.

There was also a comment in which the student said that he tried to use a chatbot in everyday life for communication, advice, and obtaining an additional opinion. However, such use “looked very silly”, so the student noted that he actively used the chatbot only for homework.

In conclusion, further research is required in this field. The problems of creative skills and critical thinking need to be studied with application of statistical analysis in order to get general results. Further studies should use statistical analysis methods to improve the reliability of the results obtained. Also, there is a need for longitudinal studies to verify the effects of AI on society.

However, we hope that our experiment contributes to the discussion. It is clear that nature of AI requires deeper understanding in the psychological characteristics of the participants, their emotional condition during the semester, predisposition to writing certain formats of texts, relationships in the team, motivation to take part in the research. AI is increasingly changing social reality. It seems relevant to study the impact of this technology on the educational process, because it is the educational sphere that determines what the society of the future will be like.

References

- Abrami et al., 2015 – Abrami, P.C., Bernard, R.M., Borokhovski, E., Waddington, D.I., Wade, C.A., Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*. 85(2): 275-18.
- Bloom et al., 1956 – Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain, David McKay Company: New York.
- Bykov et al., 2019 – Bykov, I., Gladchenko, I., Ibraeva, G., Myssaeva, K. (2019). Digital Literacy and Strategy of PR-education in Russia and Kazakhstan: Comparative Analysis. *Proceedings of the 2019 IEEE Communication Strategies in Digital Society Workshop*. Pp. 47-49. DOI: <https://doi.org/10.1109/COMSDS.2019.8709635>
- Bykov, Medvedeva, 2024 – Bykov, I.A., Medvedeva, M.V. (2024). Media Literacy and AI-technologies in Digital Communication: opportunities and Risks. *Proceedings of the 2024 Communication Strategies in Digital Society Seminar*. Pp. 21-24. DOI: <https://doi.org/10.1109/ComSDS61892.2024.10502053>
- Darwin et al., 2023 – Darwin, R.D., Mukminatien, N., Suryati, N., Laksmi, E.D., Marzuki (2023). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*. 11(1). DOI: <https://doi.org/10.1080/2331186X.2023.2290342>
- Deepfake..., 2023 – Deepfake 'news anchors' appear in pro-China footage on social media, research group say. 2023. [Electronic resource]. URL: <https://www.abc.net.au/news/2023-02-08/deepfake-news-anchors-appear-in-pro-china-footage-research/101949284>
- Essien et al., 2024 – Essien, A., Bukoye, O. T., O'Dea, X., Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*. 1(18): 865-882. DOI: <https://doi.org/10.1080/03075079.2024.2316881>
- Faqih, 2023 – Faqih, N. (2023). Developing Literacy and Critical Thinking with AI: What Students Say. *Proceeding of Annual International Conference on Islamic Education (AICIED)*. 1: 16-25. [Electronic resource]. URL: <https://prosiding.iainponorogo.ac.id/index.php/aicied/article/view/1047>
- Fedorov, Levitskaya, 2017 – Fedorov, A., Levitskaya, A. (2017). Media Education and Media Criticism in the Educational Process in Russia. *European Journal of Contemporary Education*. 6(1): 39-47.
- Fullan, 2023 – Fullan, M., Azorín, C., Harris, A., Jones, M. (2023). Artificial intelligence and school leadership: challenges, opportunities and implications. *School Leadership & Management*. 1(8). DOI: <https://doi.org/10.1080/13632434.2023.2246856>
- Guilford, 1973 – Guilford, J.P. (1973). Characteristics of Creativity. Illinois State Office of the Superintendent of Public Instruction, Springfield. Gifted Children Section. 10 p. [Electronic resource]. URL: <https://files.eric.ed.gov/fulltext/ED080171.pdf>
- Guo, Lee, 2023 – Guo, Y., Lee, D. (2023). Leveraging ChatGPT for Enhancing Critical Thinking Skills. *Journal of Chemical Education*. 100 (12): 4876-4883.
- Hapsari, Wu, 2022 – Hapsari, I.P., Wu, T.T. (2022). AI Chatbots Learning Model in English Speaking Skill: Alleviating Speaking Anxiety, Boosting Enjoyment, and Fostering Critical Thinking. In: Huang, Y.M., Cheng, S.C., Barroso, J., Sandnes, F.E. (eds) *Innovative Technologies and Learning. ICITL 2022. Lecture Notes in Computer Science*. 13449: 444-453. DOI: https://doi.org/10.1007/978-3-031-15273-3_49
- Hejres, 2022 – Hejres, S. (2022). The Impact of Artificial Intelligence on Instructional Leadership. In *Technologies, Artificial Intelligence and the Future of Learning Post-COVID-19. Studies in Computational Intelligence*. 1019: 691-711.
- Krathwohl, 2002 – Krathwohl, D.R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*. 41(4): 212-218. DOI: https://doi.org/10.1207/s15430421tip4104_2
- Nassar, 2019 – Nassar, S. (2019). Future aspects of critical thinking and AI. *Proceedings of WRFASE International Conference*. Abu Dhabi, UAE. Pp. 23-27.
- Ray, 2023 – Ray, P.P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. 3: 121-154. DOI: <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Spector, Ma, 2019 – Spector, J.M., Ma, S. (2019). Inquiry and critical thinking skills for the next generation: from artificial intelligence back to human intelligence. *Smart Learning Environment*. 6 (8). DOI: <https://doi.org/10.1186/s40561-019-0088-z>

[Van Quaquebeke, Gerpott, 2023](#) – Van Quaquebeke, N., Gerpott, F.H. (2023). The Now, New, and Next of Digital Leadership: How Artificial Intelligence (AI) will Take Over and Change Leadership as We Know It. *Journal of Leadership & Organizational Studies*. 30(3): 265-275. DOI: <https://doi.org/10.1177/15480518231181731>

[Walter, 2024](#) – Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*. 21(15). DOI: <https://doi.org/10.1186/s41239-024-00448-3>

[Yi, 2024](#) – Yi, W. (2024). Critical Thinking Pedagogics Design in an Era of ChatGPT and Other AI Tools – Shifting From Teaching “What” to Teaching “Why” and “How”. *Journal of Education and Development*. 8 (1). DOI: <https://doi.org/10.20849/jed.v8i1.1404>