https://doi.org/10.34142/2709-7986.2025.30.1.09

HOW CHATGPT SHAPES A NEW REALITY OF WRITING: IS THERE A PLACE FOR HUMANS IN AN ARTIFICIAL WORLD?

ЯК ЧАТ GPT ФОРМУЄ НОВУ РЕАЛЬНІСТЬ ПИСЬМА: ЧИ Є МІСЦЕ ЛЮДИНІ У ШТУЧНОМУ СВІТІ?

Received: 11/01/2025 **Accepted:** 12/02/2025

Veronika YURCHENKO[1], & Oleksii NALYVAIKO[2]

[1] *4th-year student, School of Psychology, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.*

✉ **E-Mail:** *veronika.ur4@gmail.com*

https://orcid.org/0009-0000-1065-9595

[2] *Ph.D. in Pedagogy, Associate Professor, Associate Professor of Pedagogy Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.*

✉ **E-Mail:** *nalyvaiko@karazin.ua*

https://orcid.org/0000-0002-7094-1047

How to Cite (APA Style):

Yurchenko, V., & Nalyvaiko, O. (2025). How ChatGPT shapes a new reality of writing: Is there a place for humans in an artificial world? *Educational Challenges, 30*(1), 138–155. https://doi.org/10.34142/2709-7986.2025.30.1.09

ABSTRACT

Aim. *This study explores the impact of Artificial Intelligence (AI)-driven text generation on authorship recognition and its implications for education. Specifically, it examines how individuals with varying levels of experience using ChatGPT perceive AI-generated texts and whether editing*

Мета дослідження. *Це дослідження вивчає вплив генерації текстів штучним інтелектом (ШІ) на розпізнавання авторства та його наслідки для освіти. Зокрема, аналізується, як люди з різним рівнем досвіду використання ChatGPT сприймають тексти, створені ШІ, і чи впливає*

© Veronika YURCHENKO, & Oleksii NALYVAIKO, 2025

How ChatGPT shapes a new reality of writing: Is there a place for humans in an artificial world? © 2025 by Veronika YURCHENKO, & Oleksii NALYVAIKO is licensed under **Attribution-NonCommercial 4.0 International**. To view a copy of the license, visit http://creativecommons.org/licenses/by-nc/4.0/

affects their ability to distinguish between human and AI-authored content. The study aims to inform strategies for integrating AI literacy into education while addressing challenges in authorship attribution.

Methodology. *A quantitative study was conducted with 85 participants, categorized based on their ChatGPT usage: no experience (n=22), occasional use (n=31), and daily use (n=32). Participants evaluated texts across five stylistic genres (literary, journalistic, scientific, philosophical, and promotional) and attempted to determine authorship (human, AI-generated, or AI-edited). The data were analyzed using statistical methods, including the Kruskal-Wallis test, ANOVA, and Pearson's correlation analysis.*

Results. *The findings indicate that participants struggled to differentiate between human and AI-generated texts, with an average accuracy of 5.48 out of 15. The ability to identify authorship varied by genre, with philosophical texts being the easiest to recognize and journalistic texts posing the greatest challenge. Experience with ChatGPT did not significantly improve recognition accuracy. Editing AI-generated texts further blurred the distinction between human and machine authorship.*

Conclusions. *The results highlight the need for educational approaches that enhance critical literacy and awareness of AI-generated content. AI tools can serve as collaborative writing aids, but their integration into learning requires ethical considerations and adapted assessment strategies. Future research should expand the sample size and investigate the effects of long-term AI exposure on writing perception.*

Keywords: *AI-generated text, ChatGPT, authorship recognition, digital literacy, artificial intelligence in education, human-AI collaboration, writing assessment.*

редагування на їхню здатність розрізняти авторство. Дослідження спрямоване на розробку стратегій інтеграції ШІ-грамотності в освітній процес та вирішення проблеми атрибуції авторства.

Методологія. *Було проведено кількісне дослідження за участю 85 респондентів, розподілених за рівнем використання ChatGPT: не використовували (n=22), використовували зрідка (n=31) і щодня (n=32). Учасники оцінювали тексти п'яти стилістичних жанрів (художній, журналістський, науковий, філософський і рекламний) та визначали їхнє авторство (написаний людиною, створений ШІ або створений ШІ з редагуванням). Дані були проаналізовані за допомогою статистичних методів, включаючи тест Крускала-Уолліса, дисперсійний аналіз (ANOVA) і кореляційний аналіз Пірсона.*

Результати. *Дослідження показало, що учасники мали труднощі з розрізненням текстів, створених людиною та ШІ, середня точність визначення авторства склала 5,48 з 15. Рівень правильного визначення авторства відрізнявся залежно від жанру: найвищі результати були у філософських текстах, а найскладніше було розпізнати журналістські. Досвід використання ChatGPT не мав значного впливу на точність розпізнавання. Редагування ШІ-текстів ще більше ускладнило їх відмінність від людських.*

Висновки. *Отримані результати підкреслюють необхідність розвитку критичної грамотності та обізнаності щодо ШІ-генерованого контенту. Інструменти штучного інтелекту можуть бути корисними у навчальному процесі, проте їх інтеграція потребує етичного підходу та адаптації методів оцінювання. Подальші дослідження мають охоплювати більші вибірки та вивчати довгостроковий вплив взаємодії з ШІ на сприйняття письмових текстів.*

Ключові слова: *ШІ-генерований текст, ChatGPT, розпізнавання авторства, цифрова грамотність, штучний інтелект в освіті, співпраця людини та ШІ, оцінювання письма.*

INTRODUCTION

In recent years, the rapid advancement of artificial intelligence (AI) technologies has significantly influenced diverse aspects of society, including academic, professional, and creative writing (OpenAI, 2022; Nalyvaiko, 2023). Among these developments, large language models (LLMs), such as ChatGPT, have emerged as transformative tools within the field of AI-driven text generation. By leveraging extensive corpora of written data and sophisticated machine learning techniques, ChatGPT can produce coherent,

contextually relevant prose that closely resembles human expression (Devlin et al., 2019). This remarkable capability has sparked widespread interest in the potential benefits of automated writing, such as enhanced productivity and innovative content creation strategies (Marcus & Davis, 2019).

At the same time, concerns have arisen regarding the extent to which AI-generated text may overshadow human authorship, creativity, and critical thinking (Bender et al., 2021). Questions persist about how these models might alter traditional writing processes, disrupt established norms of intellectual property, and reshape the labor market for writers and editors (Mateescu & Elish, 2019).

Additionally, ethical considerations regarding bias, transparency, and authenticity in AI-generated text represent critical areas for further inquiry (Floridi & Cowls, 2019). Within this emergent landscape, it becomes essential to investigate whether there is still space for uniquely human expertise and perspective in a reality increasingly shaped by automated systems. Consequently, this study aims to examine how ChatGPT contributes to a new reality of writing and to explore the enduring role of human authorship in the face of rapidly evolving AI technologies.

Research on AI-driven text generation has evolved rapidly over the last decade, building upon foundational advances in natural language processing (NLP) and machine learning. Early approaches relied on rule-based systems and simple statistical language models (Jurafsky & Martin, 2019).

More recent breakthroughs – particularly the development of large-scale transformer architectures – have significantly enhanced the fluency and coherence of machine-generated text (Devlin, Chang, Lee, & Toutanova, 2019). This transition is evident in models like GPT-3 and GPT-4, which leverage billions of parameters to predict contextually relevant tokens, creating output that closely mimics human writing (Brown et al., 2020; OpenAI, 2023).

Central to this evolution is the concept of pre-training on vast corpora of text from diverse domains. Such an approach enables AI models to capture nuanced linguistic patterns, idiomatic expressions, and domain-specific knowledge (Devlin et al., 2019). Consequently, tools like ChatGPT not only demonstrate proficiency in generating grammatically accurate sentences but also engage in more complex tasks, including summarization, translation, and even creative writing (OpenAI, 2023). While these capabilities offer unprecedented opportunities for efficiency and innovation, they also prompt questions about the boundaries between human and AI authorship (Bender et al., 2021).

The rise of increasingly sophisticated language models has led scholars to re-examine the concept of creativity and its potential overlap with AI outputs. Traditionally, creativity was viewed as a distinctly human faculty, rooted in subjective experience, emotion, and cultural context (Csikszentmihalyi, 1996). However, contemporary studies reveal that AI-driven text generation systems can produce novel ideas or stylistic flourishes that appear creative to humans (Floridi & Chiriatti, 2020; Nalyvaiko, 2023).

Despite these observations, many researchers emphasize that these AI outputs are fundamentally dependent on the data from which the models learn (Bender et al., 2021). Lacking genuine intentionality or self-awareness, AI systems rearrange patterns rather than engage in authentic generative processes. As a result, they often replicate

dominant cultural assumptions found in their training data, inadvertently reinforcing existing stereotypes or biases (Shah, Schwartz, & Hovy, 2020). This underscores the enduring importance of human oversight and ethical judgment in both the creation and application of automated writing tools (Floridi & Cowls, 2019).

The proliferation of AI-written content introduces diverse ethical and societal challenges. One primary concern involves the potential for misinformation and the difficulty that readers may face in discerning human-authored text from AI-generated prose (DiResta, 2018). As language models become more adept at emulating human stylistic nuances, the risk of using AI systems for deceptive purposes – such as generating deepfake news articles or misleading social media posts – also increases (Chesney & Citron, 2019).

Moreover, scholars highlight the importance of transparency when deploying AI-driven writing tools in academic, professional, and public contexts. Ensuring that audiences are informed about the involvement of AI in content creation is critical for maintaining trust and safeguarding intellectual property rights (Mateescu & Elish, 2019). Equally pressing are the implications for labor markets, particularly in professions tied to content creation. While some experts predict that AI systems will primarily augment rather than replace human workers, there remain concerns that widespread automation will lead to economic disruption and skill obsolescence (Brynjolfsson & McAfee, 2014).

Despite the impressive capabilities of AI-driven text generation, researchers generally concur that human expertise remains indispensable. Beyond merely generating coherent text, humans bring context-specific knowledge, cultural sensitivity, ethical discernment, and moral accountability to written communication (Floridi & Cowls, 2019). These uniquely human attributes are essential for ensuring that AI-assisted writing aligns with social, cultural, and ethical norms.

In academic and creative endeavors, collaboration between humans and AI may foster novel forms of artistic and intellectual exploration (Marcus & Davis, 2019). However, the onus is on policymakers, educators, and technology developers to shape frameworks that harness the benefits of these systems while addressing their limitations (Bender et al., 2021). By establishing robust guidelines for authorship, transparency, and accountability, society may integrate AI into the writing process without sacrificing the distinctly human aspects that underpin literary innovation and critical thought.

The aim of this study is to examine the ability of participants with varying levels of experience in using ChatGPT to recognize the authorship of texts (human or AI) and to evaluate the impact of editing artificially generated texts on their perception. The study also seeks to identify differences in the perception of texts between groups of participants who have never used ChatGPT and those who have regular experience using the tool. Additionally, the study outlines potential applications of the results in educational activities.

METHODOLOGY

The study was conducted to analyze participants' ability to determine the authorship of texts and to assess the influence of their experience with artificial intelligence on their perception of textual materials. A quantitative analysis method was applied, which included experimental evaluation of texts and statistical processing of the collected data. The study sample consisted of 85 participants, divided according to their level of

ChatGPT usage: 22 individuals (25.9%) had no experience using AI, 31 individuals (36.5%) used ChatGPT occasionally (a few times a month), and 32 individuals (37.6%) used AI daily.

The study utilized five blocks of texts from different stylistic categories:

- Literary (Kidruk, 2017),
- Journalistic (Kapustin, 2021),
- Scientific (Nalyvaiko & Chornous, 2017),
- Philosophical (Pascal, 1901),
- Promotional (GlobalLogic, 2020).

Each block contained three text variants:

1. Human-written text – created by a human;
2. AI-generated text – produced by ChatGPT without modifications;
3. AI-generated text with edits – a ChatGPT-generated text edited by providing additional instructions to improve its alignment with human-like style.

For generating the AI texts, specially formulated prompts were used, requesting the creation of a similar text in the corresponding style. Initially, the generated texts were included in the survey in their unmodified form, as well as in an edited version where the AI made adjustments based on additional instructions.

The survey was conducted via Google Forms. Participants received a random set of texts and were asked to determine their authorship by choosing one of three options: “human”, “AI”, or “AI with edits”. Participants also answered additional questions aimed at identifying the criteria they used to determine authorship and assessed changes in their trust in the text upon detecting “AI markers”.

Statistical data processing included comparative analysis of the results between participant groups using the Kruskal-Wallis test, Levene's test, analysis of variance (ANOVA), and Pearson's correlation analysis. The relationship between the level of AI experience and the accuracy of determining text authorship was also examined.

The study adhered to ethical standards: all participants provided informed consent, their responses were anonymized, and the data were used exclusively for scientific purposes.

The results of the experiment allow for an evaluation of the impact of AI on the perception of texts and the formation of trust in them, as well as the identification of key factors influencing the recognition of artificially generated texts.

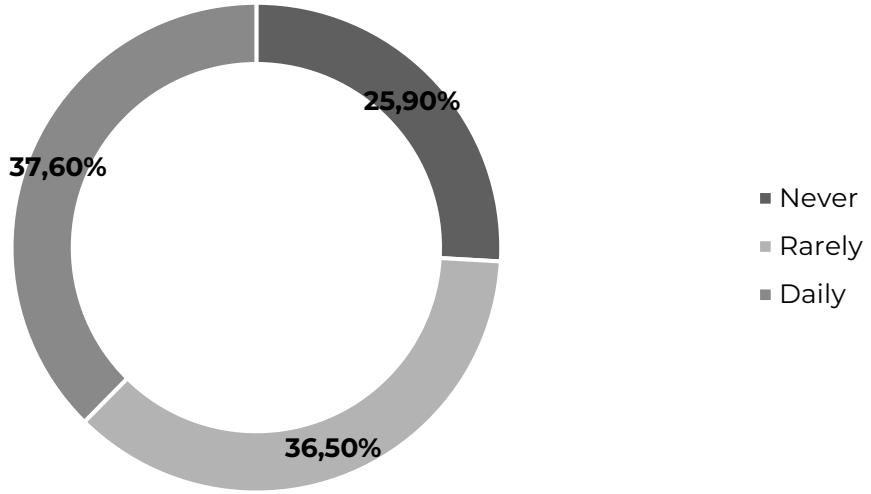
RESULTS

Demographic Characteristics of the Sample

The study involved 85 respondents. By gender: 85% were women and 15% were men. The age range of participants varied from 16 to 57 years, with the average age being 26 years. Regarding professional activity, 57% of participants were students, 30% were employed, 7% were unemployed, and 5% were school students. The primary objective of the study was to assess the respondents' ability to determine the authorship of texts, so a key aspect of the analysis was considering the level of ChatGPT usage (Fig. 1).

Figure 1

Frequency of ChatGPT usage



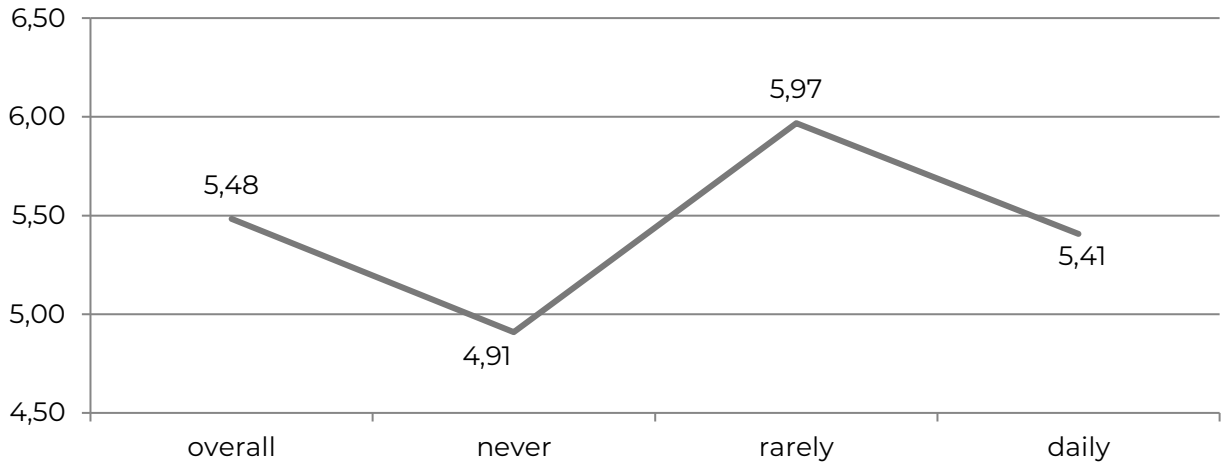
Quantitative Analysis of the Results

The analysis of the overall results showed that the average number of correct answers among all respondents was 5.48 out of 15 possible (Fig. 2). This indicates the general difficulty in recognizing the authorship of texts, which may be due to the increasing similarity between texts created by humans and artificial intelligence. However, a more detailed study revealed significant differences in the ability to recognize authorship depending on the level of AI usage.

For instance, participants who had never used ChatGPT correctly identified an average of only 4.91 texts, while those who used AI occasionally (a few times a month) demonstrated the best result – 5.97 correct answers. Participants who used ChatGPT daily had a slightly lower score – 5.41, suggesting a possible influence of familiarity with the stylistic features of AI, which complicates their recognition.

Figure 2

Average number of correctly identified texts based on ChatGPT usage frequency



When examining the results by text styles, it was found that the level of correct authorship identification significantly varies depending on the genre of the texts (Table 1).

Table 3

The level of correct determination of text authorship depending on the style and frequency of use of ChatGPT (H – human text, ED – AI with edits, AWC – AI without changes)

		H	ED	AWC	H	ED	AWC	H	ED	AWC	H	ED	AWC	H	ED	AWC
Gen.	85	33%	32%	39%	33%	29%	46%	28%	35%	44%	41%	40%	46%	32%	38%	33%
Never	22	32%	36%	45%	23%	27%	36%	18%	36%	32%	41%	32%	45%	27%	32%	27%
Rarely	31	42%	32%	42%	39%	29%	52%	42%	39%	48%	39%	35%	39%	35%	48%	35%
Daily	32	25%	28%	31%	34%	31%	47%	22%	31%	47%	44%	50%	53%	31%	31%	34%
		literary			**journalistic**			**scientific**			**philosophical**			**promotional**		

The distribution of correct answers among all respondents indicates that the highest level of recognition is observed in the philosophical style (41% correct answers for human-written texts), while journalistic texts proved to be the most challenging for authorship identification (only 28% correctly identified human-written texts). This can be explained by the fact that philosophical texts have a more specific style, which is harder for AI to accurately reproduce, whereas journalistic writing is characterized by a general structuredness that allows AI to effectively mimic it.

In the literary style, respondents most often made mistakes by identifying AI-generated texts as human-written. The overall level of correct recognition of human-written texts in this category was 33%, indicating that text generation algorithms adapt well to artistic stylistics. At the same time, respondents who use AI daily had the lowest level of correct answers (25%), suggesting a decrease in critical evaluation due to familiarity with artificially generated content.

In the scientific style, the level of correct answers for human-written texts was the lowest – 28%. This supports the hypothesis that AI effectively imitates scientific discourse by using a clear structure, formal expressions, and technical vocabulary. The best results in this category were demonstrated by respondents who occasionally use AI (39% correct answers).

Advertising texts showed an average level of recognition accuracy, but a high rate of errors was noted among those who had never used AI. This may be due to the fact that advertising style often follows clearly defined templates, which are easy for neural networks to replicate, making authorship identification more difficult.

Evaluation of Criteria for Determining Text Authorship

To analyze the mechanisms of text perception by respondents, a survey was conducted to identify the factors influencing the determination of text authorship.

Fig. 3 illustrates the most significant factors that influenced respondents' decisions regarding text authorship. The analysis of responses indicates that the defining criteria were the style of the text, followed slightly by grammatical and lexical correctness, while the logic of presentation and the level of emotionality were equally decisive for more than half of the respondents. This confirms the assumption that structural and stylistic features are key in the process of authorship identification.

Figure 3

Factors influencing text authorship identification

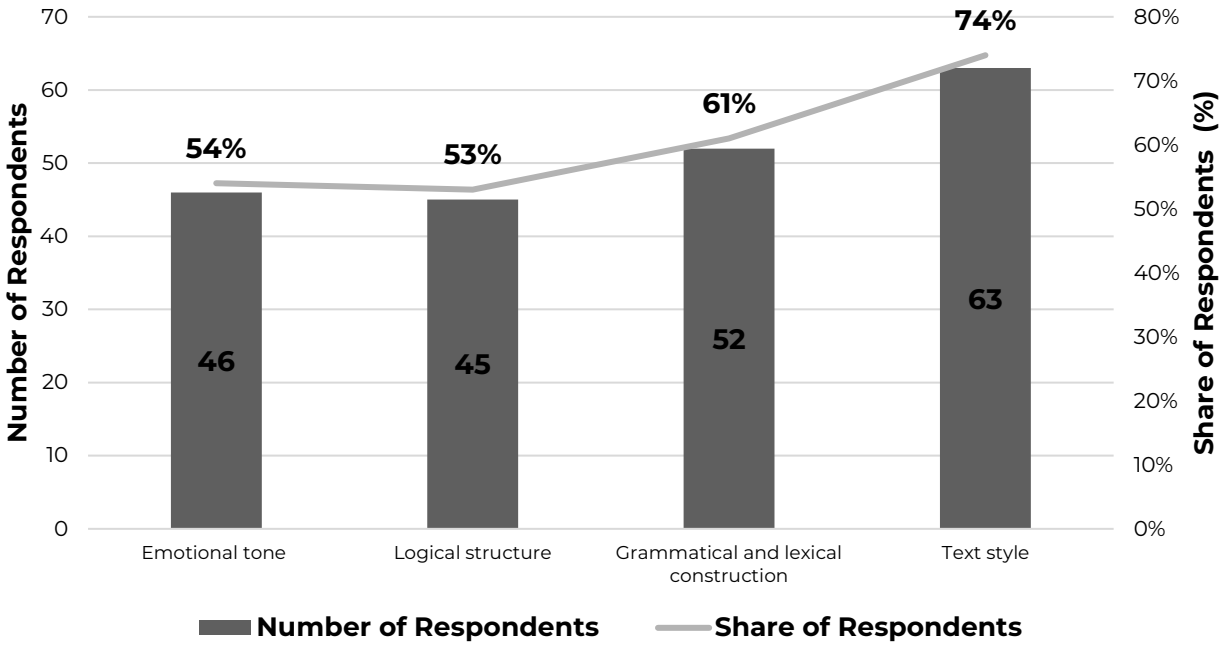
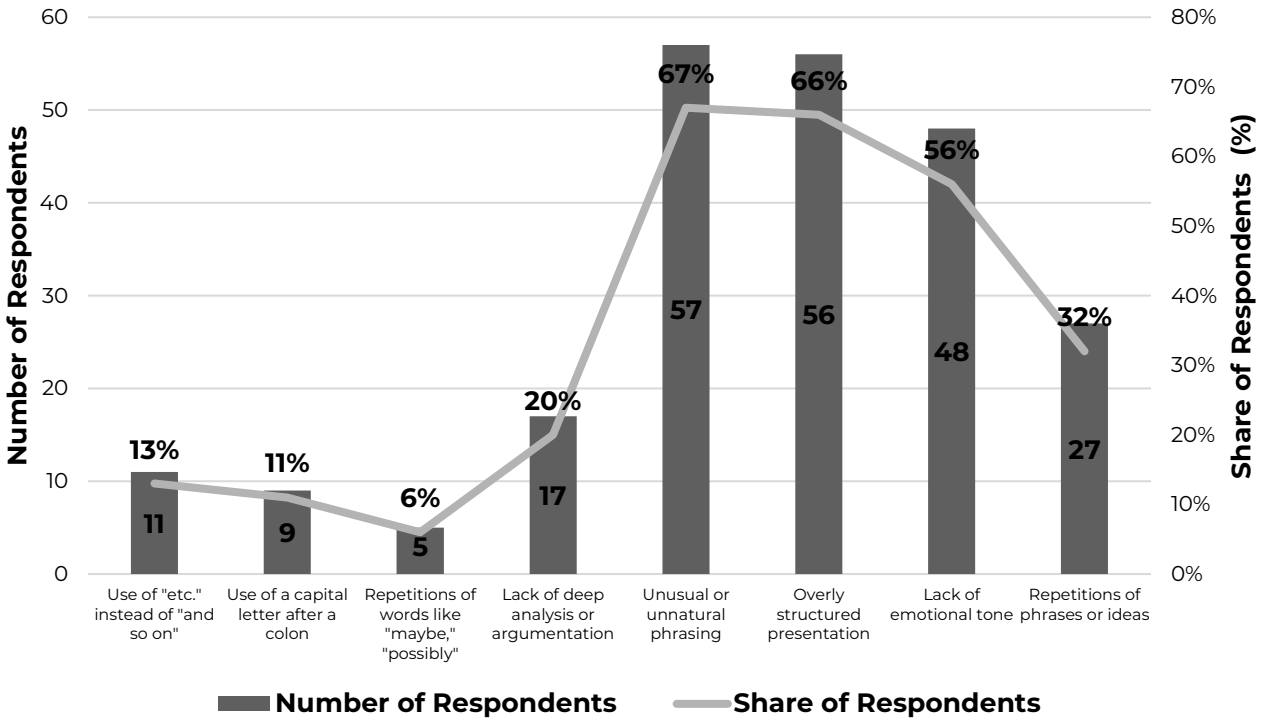


Fig. 4 shows the most common features that raised suspicions about the artificial origin of the text were the lack of emotional nuance, excessive structure, repetitive phrases, and the absence of in-depth analysis. Some formal characteristics, such as the use of formulaic expressions or violations of grammatical rules, also contributed to identifing the text as AI-generated.

Figure 4

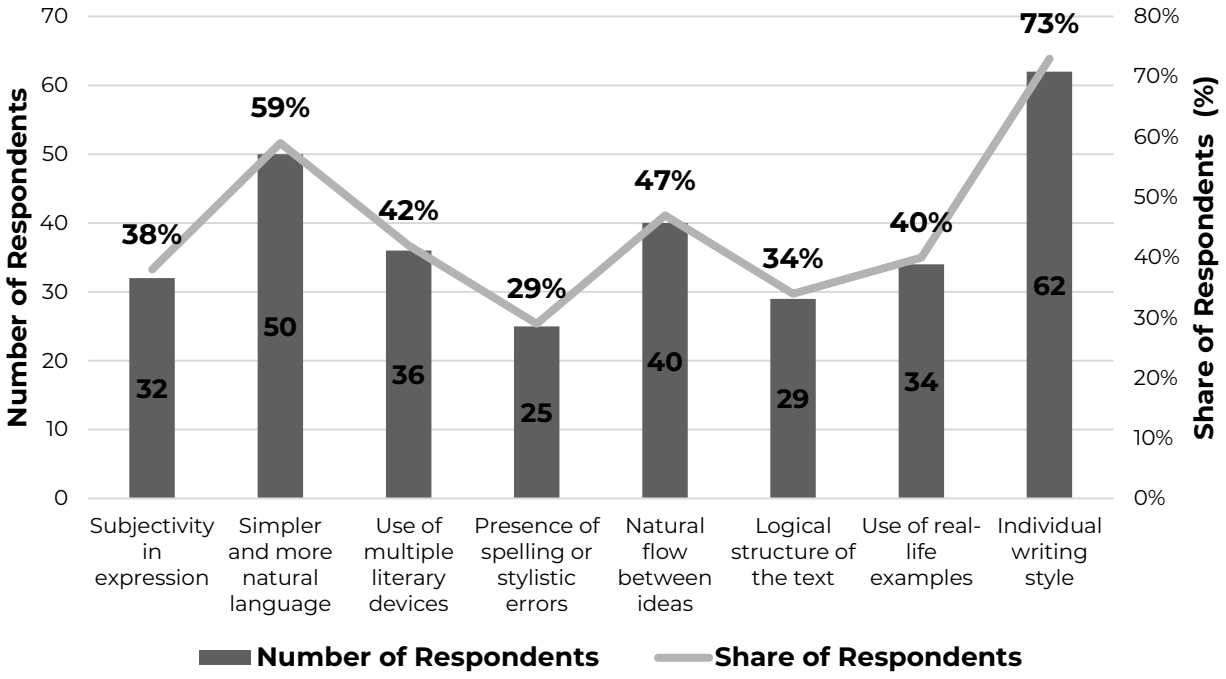
Key characteristics perceived as artificial



According to Fig. 5, the most characteristic features of human-written texts, as perceived by respondents, were individual style, simple language of presentation, natural transitions between ideas, and the use of literary tropes. The least influential factors were spelling and stylistic errors, which were likely perceived as evidence of the absence of algorithmic text processing.

Figure 5

Key characteristics perceived as human-like



Within the framework of the study, respondents were asked to determine the authorship of texts and explain their assumptions. The analysis of open-ended responses allows us to outline the main patterns in the perception of human-written and AI-generated texts, as well as the difficulties encountered by the participants in the experiment.

1. Criteria Associated with AI-Generated Texts

The respondents' answers demonstrate that the main features of texts perceived as AI-generated include structuredness, lack of emotional tone, and excessive formality. For example, one respondent noted: *"A cold text, without emotional coloring"*. Another emphasized that AI texts often have *"an overly structured presentation with words that people do not use in everyday life"*. Some participants pointed out AI's tendency to use complex grammatical constructions rarely found in natural speech: *"Rather complicated structures that people usually do not use nowadays"*.

In addition to syntactic features, respondents paid attention to unnatural language patterns. For instance, specific terms atypical for colloquial language were identified: *"I looked for errors – for example, one text used the phrase 'mobile gadgets.' I've never heard a person use that, so I'm sure it was AI"*. AI's tendency toward generalization and avoidance of subjectivity was also noted: *"AI strives to generalize and lacks a personal opinion"*.

Interestingly, some respondents suggested that excessive emotionality could also signal artificial origin: *"Overly emotional language and an abundance of epithets in every literary sentence looked suspiciously artificial"*. This indicates that, despite the common perception of AI texts as "dry," the exaggeration of certain stylistic techniques may also raise doubts about their authorship.

2. Features of Human-Written Texts

The respondents' answers indicate that texts identified as human-written often contained individual style, emotional tone, and natural variability in speech. For example, one participant stated: *"Simplicity in writing, greater emotionality, and reader engagement"*. Others noted the use of stylistic devices typical of literary or journalistic language: "*Many adjectives, aphorisms, metaphors"*.

Some respondents identified human texts based on syntactic errors or stylistic inconsistencies: "*In some parts, punctuation was missing, so it was likely written by a human"*. Others emphasized natural flow: "*Simpler language and logical progression of ideas"*.

Notably, some participants highlighted subjectivity and inconsistency as markers of human writing: *"A balance between emotionality and dryness in literary texts. Humans often write as they would speak"*. This aligns with the concept of cognitive variability inherent in natural human language.

3. Challenges in Determining Authorship

Despite the respondents' ability to identify certain markers, many found the task difficult. One participant stated plainly: *"It's actually hard to tell"*. Others noted similarities between texts: *"Some texts were too similar, making it hard to distinguish 'authentic' ones"*.

Some remarked on AI's ability to mimic human style convincingly: *"AI skillfully copies human writing styles, making it hard to differentiate"*. One respondent explained: *"AI uses many epithets and tries to make the text as vivid and structured as possible"*.

Others acknowledged cognitive bias influenced their judgment: "*The research task itself triggered a biased attitude toward every text and heightened suspicion. If I had read the text first and then been asked about authorship, accuracy might have improved"*. This suggests that the task's framing may have skewed results by prompting respondents to seek differences even where minimal.

Participants also struggled to distinguish AI-generated texts from AI-edited texts: *"In my opinion, literary texts are harder to identify because AI can also emulate emotionality. However, it might overdo it"*.

Methods of Mathematical Statistics

To analyze respondents' perception of texts from different genres, methods of mathematical statistics were applied. The main criteria included Levene's test, Kruskal-Wallis test, analysis of variance (ANOVA), and Pearson's correlation analysis. This allowed for an assessment of how text authorship recognition varied depending on the frequency of ChatGPT usage and identified relationships between genres.

1. Levene's Test (Table 2)

Levene's test was used to check the homogeneity of variances among respondent groups with different levels of ChatGPT usage. Since the overall p-value of Levene's test is 0.025, indicating heterogeneity of variances, the Kruskal-Wallis test is the primary method for testing group differences. However, for individual genres where $p > 0.05$, using ANOVA is appropriate.

Table 2

Assessment of variance homogeneity among respondent groups (Levene's Test)

Genre	Levene's Test	p-value
Overall	3.865	0.025
Literary	2.760	0.069
Journalistic	0.892	0.414
Scientific	1.859	0.162
Philosophical	1.119	0.332
Promotional	0.146	0.864

2. Kruskal-Wallis Test (Table 3)

This test was used to verify the hypothesis of equal response distributions across respondent groups (never, rarely, daily). The results showed that the response distributions did not exhibit statistically significant differences between groups ($p > 0.05$). This suggests that the frequency of ChatGPT usage does not significantly impact the ability to correctly identify text

Table 3

Hypothesis testing results on differences in text recognition accuracy (Kruskal-Wallis Test)

Genre	H-value	p-value	Decision
Overall	2.281	0.320	Null hypothesis accepted
Literary	2.885	0.236	Null hypothesis accepted
Journalistic	2.152	0.341	Null hypothesis accepted
Scientific	2.092	0.351	Null hypothesis accepted
Philosophical	1.526	0.466	Null hypothesis accepted
Promotional	2.688	0.261	Null hypothesis accepted

3. Analysis of Variance (ANOVA) (Table 4)

To assess whether the frequency of ChatGPT usage affects the ability to correctly determine text authorship in different genres, ANOVA was conducted (Table 4). The results showed that neither the overall effect nor the analysis of individual genres provided evidence for significant differences in text authorship identification accuracy depending on AI usage levels. All p-values exceeded 0.05, indicating no statistically significant effects. These results suggest that the level of experience in interacting with

ChatGPT does not influence respondents' ability to correctly identify text authorship across different.

Table 4

ANOVA results on the effect of ChatGPT usage frequency

Genre	F-value	p-value
Generally	1,240	0,295
Literary	0,912	0,406
Journalistic	0,619	0,541
Scientific	1,316	0,274
Philosophical	0,788	0,458
Promotional	0,738	0,481

4. Pearson's Correlation Analysis (Table 5)

The correlation analysis between text genres revealed generally low relationships between styles. The highest correlation coefficient was observed between the scientific and journalistic styles (r = -0.62), indicating some structural differences in these texts. Other correlations were minor, suggesting weak or no interdependencies between genres.

Table 5

Pearson's correlation analysis between text genres

Genre	Literary	Journalistic	Scientific	Philosophical	Promotional
Literary	–	-0.045	0.126	0.074	-0.024
Journalistic	-0.045	–	-0.62	-0.085	-0.097
Scientific	0.126	-0.62	–	0.140	0.045
Philosophical	0.073	-0.085	0.140	–	0.066

Thus, the analysis confirmed that the frequency of ChatGPT usage does not have a significant impact on respondents' ability to recognize text authorship. Correlation analysis indicated that the highest stylistic similarity was observed between scientific and philosophical texts, while the most notable differences were between journalistic and scientific texts. These findings may be useful for further research on the impact of artificial intelligence on text perception.

5. Regression Analysis

To determine the impact of AI usage frequency, response accuracy, and confidence in responses on the level of trust in AI-generated texts, multiple linear regression analysis was conducted. Trust served as the dependent variable, while AI usage frequency, response accuracy, and confidence in responses were predictors.

The results of the ANOVA for the regression model (Table 6) indicate that the model is not statistically significant overall, suggesting that the selected factors do not have a substantial impact on trust in texts.

Table 6

ANOVA results for assessing the impact of AI usage frequency, accuracy, and confidence on trust in texts.

Model	Sum of Squares	df	Mean Square	F-value	p-value
Regression	2.400	3	0.800	1.953	0.128b
Residual	33.176	81	0.410		
Total	35.576	84			

The regression coefficient estimates (Table 7) showed that confidence in responses had the highest standardized coefficient (β), indicating a tendency toward a weak positive correlation between confidence and trust, although this indicator does not reach statistical significance. AI usage frequency had a negative but insignificant effect, and the overall response accuracy level had almost no impact on trust.

Table 7

Regression analysis of the impact of AI usage frequency, accuracy, and confidence on trust in texts.

Predictor	B	Standard Error	β	t	p-value
(Constant)	1.675	0.359	–	4.669	<0.001
Confidence	0.123	0.072	0.185	1.698	0.093
Frequency	-0.122	0.089	-0.149	-1.375	0.173
Accuracy	0.025	0.029	0.093	0.860	0.392

Thus, the obtained results did not confirm a statistically significant impact of AI usage frequency, response accuracy, or response confidence on trust in texts. However, the identified tendency for a positive relationship between response confidence and trust may indicate individual differences in how respondents perceive textual information.

DISCUSSION

The integration of artificial intelligence into the realm of writing is reshaping the way individuals engage with text, posing both challenges and opportunities for education. The findings of this study provide a deeper understanding of the interplay between human and AI-generated writing, shedding light on how these evolving dynamics influence learning activities and academic practices.

More than ever, students and educators must navigate a landscape where distinguishing between human and machine-authored content is not always straightforward. This section explores the educational benefits of the study, its broader societal relevance, and the potential directions for future research.

One of the most striking findings of this research is the difficulty participants encountered in distinguishing AI-generated texts from human-authored ones. This growing ambiguity highlights an urgent need to integrate critical literacy training into educational curricula. Traditional notions of textual analysis – once focused on argumentation, structure, and rhetorical devices – must now expand to include the

ability to identify subtle markers of AI authorship, such as excessive structural uniformity, limited emotional depth, or unnatural phrase constructions.

Educators can harness these insights by designing instructional activities that challenge students to evaluate the authenticity of texts, fostering a habit of careful scrutiny. By doing so, learners develop not only a keen eye for textual patterns but also a more sophisticated understanding of how AI tools operate. In a world increasingly shaped by synthetic content, the ability to assess and verify information is becoming a fundamental academic skill.

The study's findings suggest that AI-generated texts – especially when refined through human intervention – can closely mimic human writing. This opens up new possibilities for using AI as a collaborative tool rather than a mere content generator. ChatGPT and similar models can support students in multiple ways:

- A) Writing Support (AI can serve as a brainstorming partner, helping students structure their ideas before they refine and personalize the final text);
- B) Creativity Enhancement (by providing stylistic variations and alternative phrasing, AI can encourage students to experiment with different literary techniques);
- C) Automated Feedback (AI-driven systems can assist in reviewing drafts, offering instant feedback on grammar, coherence, and argumentation, allowing educators to focus on more nuanced aspects of writing instruction).

These applications do not replace the need for human oversight but rather extend the possibilities for personalized learning, encouraging students to critically engage with both their own writing and AI-generated suggestions.

The study also raises pressing questions about assessment strategies in the age of AI-assisted writing. If AI-generated texts can closely resemble human-authored ones, traditional assessment methods may need to evolve. Educators might consider shifting their focus from final outputs to process-oriented evaluations that track a student's progression from initial drafts to final versions.

Potential adaptations could include: encouraging students to annotate AI-generated text with their revisions, demonstrating their active role in shaping the output; assignments that require students to engage with AI responsibly, such as comparing unedited AI responses with human-edited versions; tasks that emphasize interpretation, critical thinking, and cultural awareness – areas where AI still lacks depth – ensuring that human intellectual contributions remain central to the writing process. These strategies align with broader efforts to cultivate not just technical proficiency but also ethical responsibility in AI-assisted learning.

Beyond the classroom, the findings of this study reflect a broader shift in how society perceives authorship, creativity, and intellectual labor. As AI-driven tools become increasingly integrated into professional and creative domains, there is an urgent need for institutional frameworks that address the evolving nature of authorship. Several key areas emerge as critical considerations. Firstly, institutions must clarify who holds authorship rights when AI contributes significantly to a written work.

Secondly, policies should mandate clear attribution when AI is used in academic, journalistic, or professional writing to maintain ethical standards. Thirdly future professionals must be equipped with the skills to collaborate with AI while preserving uniquely human competencies such as critical thinking, ethical judgment, and emotional intelligence.

These challenges are not insurmountable but require a proactive approach to integrating AI in a way that enhances, rather than undermines, human expertise.

The study has several limitations that affect the generalizability of the obtained results. Firstly, the small sample size reduces the representativeness of the findings and complicates their extrapolation to a broader population. Additionally, the lack of clear criteria for classifying the level of ChatGPT usage, particularly in defining the frequency and intensity of interaction, led to variability in grouping respondents, which impacted the accuracy of the statistical analysis.

A significant limitation is also the cognitive bias of respondents regarding the authorship of texts, as the awareness of the presence of texts generated by artificial intelligence influenced their evaluation. As one participant noted: *"The research task immediately provoked a biased attitude towards each text and increased suspicion."* This indicates that respondents actively searched for markers of artificiality, even in cases where they were not obvious. Thus, the results may reflect not only actual differences between the texts but also the influence of expectations and stereotypes regarding AI-generated content.

A separate methodological limitation is that AI-generated texts are practically never left without human intervention. Any query to a language model is a form of editing, as it shapes the content and style of the response. This complicates the establishment of a clear boundary between texts created by humans and those generated by AI. Further research should focus on increasing the sample size, standardizing evaluation criteria, and considering human-AI interaction in real-world usage scenarios.

CONCLUSIONS

This study underscores the transformative potential of AI in education, revealing both its disruptive and enabling roles. The challenge is not to resist AI's integration into learning but to develop frameworks that ensure its responsible and ethical use (Kostikova et al, 2024; Chetveryk & Veretiuk, 2024).

By fostering AI literacy (Roozafzai, 2024), adapting assessment strategies, and maintaining a strong emphasis on human creativity and judgment, educators can equip students with the skills they need to navigate the evolving landscape of writing. Ultimately, the goal is not to diminish human ingenuity but to amplify it, ensuring that education remains a space where critical thought, originality, and ethical responsibility continue to thrive.

Further research should focus on expanding the sample and involving respondents with different levels of experience in using generative language models, which will allow us to more accurately determine the factors that influence the perception of AI-generated texts. An important direction is to study the influence of stylistic and structural features of the text on its identification as human or artificial intelligence-created.

The use of experimental approaches with variable changes in texts (emotionality, complexity, logical structure) is promising, which will allow us to identify key markers by which respondents distinguish texts (Rezvan & Krokhmal, 2022). Special attention should be paid to studying effective methods for explaining the process of AI text generation, which will probably increase users' trust in such systems.

In addition, further research should take into account the role of AI-generated texts in various communicative environments, in particular in journalism, education, and professional communication. Identifying optimal strategies for adapting such texts will facilitate their effective use and integration into areas where automated content creation is becoming increasingly common.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interest regarding the publication of this paper.

FUNDING

The authors declare that this study received no specific financial support.

REFERENCES

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems, 33,* 1877–1901. https://doi.org/10.48550/arXiv.2005.14165

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton. https://psycnet.apa.org/record/2014-07087-000

Chesney, R., & Citron, D. K. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review, 107*(6), 1753–1820. https://doi.org/10.15779/Z38RV0D15J

Chetveryk, V., & Veretiuk, T. (2024). The potential of artificial intelligence tools in foreign language teaching and learning. In *Modern Challenges and Achievements of the Scientific Community of the 21st Century: Proceedings of the XLIII International Scientific and Practical Conference* (pp. 139–142). Narva, Estonia. https://dspace.hnpu.edu.ua/handle/123456789/16526

Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. HarperCollins.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational*

Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

DiResta, R. (2018). Computational Propaganda: If You Make It Trend, You Make It True. *The Yale Review, 106*(4), 12–29. https://doi.org/10.1353/tyr.2018.0030

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines, 30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.8cd550d1

GlobalLogic (2020, November 12). *Media Support TechTalk in Lviv*. DOU. https://dou.ua/calendar/34915/

Jurafsky, D., & Martin, J. H. (2019). Logistic Regression. In *Speech and Language Processing* (3rd ed.). Draft. Stanford University. https://web.stanford.edu/~jurafsky/slp3/slides/5_LR_Apr_7_2021.pdf

Kapustin, O. (2021, May 8). *Pandemichna uhoda. Shans dlia svitu ta Ukrainy* [Pandemic Agreement: A Chance for the World and Ukraine]. Ukrayinska Pravda. https://www.pravda.com.ua/columns/2021/05/8/7292793/ [in Ukrainian].

Kidruk, M. (2017). *Ne ozyraysya i movchy* [Don't Look Back and Be Silent]. Kharkiv, Ukraine: Knyzhkovyi Klub “Klub Simeinoho Dozvillia” [in Ukrainian].

Kostikova, I., Holubnycha, L., Besarab, T., Moshynska, O., Moroz, T., & Shamaieva, I. (2024). Chat GPT for Professional English Course Development. *International Journal of Interactive Mobile Technologies (iJIM), 18*(02), pp. 68–81. https://doi.org/10.3991/ijim.v18i02.46623

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon.

Mateescu, A., & Elish, M. C. (2019). *AI in Context: The Labor of Integrating New Systems and Maintaining Old Ones*. New York: Data & Society Research Institute. https://datasociety.net/wp-content/uploads/2019/01/DataandSociety_AIinContext.pdf

Nalyvaiko, O. O. (2023). Perspektyvy vykorystannia neiromerezh u vyshchii osviti Ukrainy [Prospects of using neural networks in higher education of Ukraine]. *Information Technologies and Learning Tools*, *97*(5), 1–17. https://doi.org/10.33407/itlt.v97i5.5322 [in Ukrainian].

Nalyvaiko, O. O., & Chornous, N. A. (2017). Vyznachennia suti poniattia «mobilnoho navchannia» [Determination essence of the concept «mobile learning»]. *Scientific Notes of the Pedagogical Department*, *40*, 118–124. https://periodicals.karazin.ua/pedagogy/article/view/14863 [in Ukrainian].

OpenAI. (2022, November 30). *ChatGPT: Optimizing Language Models for Dialogue*. https://openai.com/blog/chatgpt

OpenAI. (2023, March 14). *GPT-4 Technical Report*. https://openai.com/research/gpt-4

Pascal, B. (1901). *The thoughts of Blaise Pascal* (C. Kegan Paul, Trans.). George Bell and Sons.

Rezvan, O., & Krokhmal, A. (2022). Peculiarities of Pedagogical Internship Organization for Future Specialists of Foreign Philology in Blended Learning. *Educational Challenges, 27*(2), 169–184. https://doi.org/10.34142/2709-7986.2022.27.2.12

Roozafzai, Z. S. (2024). Artificial Intelligence Assistance and Cognitive Abilities: Harnessing AI-Assisted Heuristic Methods for Transitioning from Critical to Creative Thinking in English Language Learning. *Educational Challenges, 29*(2), 339–361. https://doi.org/10.34142/2709-7986.2024.29.2.23

Shah, D. S., Schwartz, H. A., & Hovy, D. (2020). Predictive biases in natural language processing models: A conceptual framework and overview. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5248–5264). Online: Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.468