

UDC: 37.015.3:811.111:004.738.5

Inna Zaiarna

PhD in Education, Associate Professor,
Associate Professor at Foreign Languages Department,
Educational and Scientific Institute of Law, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine;
Leading Researcher at Comparative Studies Department for Information and Education Innovations, Institute for Digitalisation of Education of the National Academy of Educational Sciences of Ukraine, Kyiv, Ukraine
ORCID ID 0000-0002-9464-096X
izaiarna@gmail.com

Olena Zhyhadlo

PhD in Linguistics, Associate Professor,
Associate Professor at Foreign Languages Department,
Educational and Scientific Institute of Law, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID ID 0000-0002-1605-7242
olena.zhyhadlo@gmail.com

Olha Dunaievskaya

PhD in Linguistics, Associate Professor,
Associate Professor at Foreign Languages Department,
Educational and Scientific Institute of Law, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
ORCID ID 0000-0002-0915-428X
o.dunayevska@gmail.com

CHATGPT IN FOREIGN LANGUAGE TEACHING AND ASSESSMENT: EXPLORING EFL INSTRUCTORS' EXPERIENCE

Abstract. The paper presents the results of the first stage of the research into the specifics of ChatGPT application in assessment development for English as a Foreign Language (EFL) courses. The article aims to investigate the perceptions of English language teachers regarding the utilization of ChatGPT in their teaching and assessment practices obtained through a survey in order to establish the degree of their familiarity with the possibilities of ChatGPT for teaching a foreign language and the mechanics of AI-generated test tasks design. The participants are mostly experienced English language instructors, affiliated with higher education institutions in Ukraine, the EU and the USA. Findings reveal varying levels of respondents' confidence in ChatGPT's functionalities. A considerable proportion of educators appeared to be either unfamiliar with ChatGPT or possess limited knowledge of its capabilities. Hesitancy in implementing ChatGPT in teaching practice primarily stems from a lack of guidance as to its effective utilization closely followed by academic integrity concerns. However, educators who have employed ChatGPT recognize its potential as an assistant in lesson planning, content creation, task design, and assessment, especially for vocabulary, grammar, and writing. Despite acknowledging the tool's benefits, respondents express reservations about ChatGPT-generated tasks' accuracy for assessment purposes. Challenges include the need for meticulous proofreading, issues with biased information, and insufficient task complexity. Nonetheless, overall satisfaction levels among educators implementing ChatGPT range from moderate to high, substantiating its value in foreign language teaching. Thus the survey stage of the research revealed the growing interest in ChatGPT's potential for foreign language teaching and assessment among Ukrainian and international EFL and ESP instructors and underscored the need for guidelines and recommendations to create effective ChatGPT-generated test tasks. Future research will develop criteria for evaluating ChatGPT-generated tasks, provide prompt design guidelines, and formulate ethical policies for AI in assessment, addressing educators' concerns and promoting AI integration in academia.

Keywords: artificial intelligence (AI); ChatGPT; testing and assessment; foreign language competence; survey.

1. INTRODUCTION

Statement of the problem. EU recent policies, which are in line with similar efforts in Ukrainian educational space, aim at implementing digitalization into studying as well as teaching processes. These endeavours require academia in general and educators in particular to advance considerably in terms of the newest technologies and artificial intelligence (AI) tools. Generative AI has already been changing educational landscapes with AI-powered learning systems being increasingly introduced to secondary and tertiary education. Growing popularity of AI in education entails considering a number of questions such as understanding the role of a human and a machine in education, correlation of pedagogical innovations and the level of educators' and students' technical awareness, ethical concerns related to data ethics, information and academic integrity. These are just a few issues highlighted in scientific publications and social media nowadays. Various contexts in which the influence of AI on education appears the most significant can be grouped into three major dimensions – learning for AI, learning about AI and learning with AI [1]. Learning for AI highlights a need to understand and address concerns and risks associated with AI ethics, bias and fake information generated by AI-powered learning systems, while learning about AI involves preparing future AI specialists and training educators and students on mathematics, data analysis and coding.

Technical advancements in the field of Generative AI resulted in the emergence of various AI tools. Learning with AI implies integrating AI applications, which are now available as commercial products, into institutional administration, teaching and learning. Institution-facing applications of AI may be used to manage an educational institution's finance, seeking grants or monitoring students' progress. In teaching and learning the use of AI-powered writing tools and chatbots proved to be helpful both for educators and students. ChatGPT, a platform that enables human-like conversations with advanced AI, whose GPT-3.5 model can be accessed for free, is certainly among the latter. Its advanced natural language processing capabilities and an ability to generate coherent responses and handle various language tasks make it a versatile assistant for language teachers, providing support in various aspects of language training. As integration of ChatGPT into EFL and ESP classroom through task design, administering assessment as well as feedback generation seems undisputable, its potential as a teaching and learning assistant has already been exposed to keen scrutiny of scholars. However, language instructors' experience with ChatGPT as a teaching tool, and challenges and limitations encountered by the former when incorporating ChatGPT into their teaching and assessment practices seems insufficiently explored by educational researchers. Identifying these challenges can facilitate devising strategies to overcome them in teaching and learning a foreign language, and pinpoint areas for improvement in ChatGPT's functionality.

Analysis of recent research and publications. The request for advancements in Ukrainian higher education sector in general and EFL and ESP teaching in particular derives greatly from EU standards. One of the modern tendencies in EU is digitalization. In 2021 the EU launched Europe Digital Decade, which will last up to year 2030. Among its numerous objectives is the statement “*Everyone can participate in digital opportunities / no one is left behind*” which, on the one hand, encourages the representative of every profession to realise their digital tools potential. On the other hand, it substantiates the right of the academia teachers to use AI tools in order to release the tension of educator's heavy workload. The application of a Chatbot might be an option.

There are at least 10 popular Chatbots, which own a potential to assist educators in managing the challenges faced while handling various tasks, such as collecting the information for the lecture or seminar or scientific report, generating the samples of tests and questionnaires under time and environment constraints by reducing or eliminating time spent on library search engines (often pre-paid ones), which can be particularly useful in times of power blackouts and potential threats of air attacks. Among the most famous AI Chatbots are: ChatGPT, Jasper Chat, Chatsonic, Gemini, Chat by Copy.ai, Claude, Perplexity AI, Microsoft Copilot, You.com,

Character AI. The choice of the most fitting Chatbot, however, is predetermined by its free-of-charge accessibility, since not all educators are able to pay fees for such an assistance.

ChatGPT, which stands for Generative Pre-trained Transformer, is an original AI Chatbot, the type of generative, multimodal large language model used by OpenAI. It enables the biggest range of manipulations, namely it is able to understand and generate images, code, files, and texts with the help of back-and-forth conversation style. It has a free version and can produce conversations concerning any industry, topic or interest. The chat even passed the Bar Exam, which might serve as a substantial argument as far as the credibility of ChatGPT-generated text is concerned. The first version of open-access ChatGPT was presented in November 2022, and since then it was described as a non-analog AI platform [2].

Jasper Chat is a chatbot specifically designed for business professionals and writing teams. It is an excellent tool for marketing tasks, such as writing blogs and filling in the websites.

Chatsonic is a Chatbot for independent marketers. It uses GPT-4 as a foundation and has a limited amount of 10, 000 words monthly as its free plan.

Gemini is Google's advanced conversational chatbot provided with support via Google AI. It is advisable for those who use a lot of Google products day to day. Using this Chatbot can be somewhat isolated due to referring entirely to Google.

Chat by Copy.ai is a tool matching the requirements put by sales teams. It applies AI to help sales teams sell. It has been designed for sales and marketing.

Claude is an AI chatbot created primarily for analysing long text documents consisting of past messages in the conversations and uploaded documents by summarising content with spotting specific information. This AI tool turned out to be of particular assistance for researchers as it was reported to be effective in understanding long articles, answering questions, solving problems, and writing. Claude is available only to Beta users in selected countries.

Perplexity AI is a chatbot used for making simple searches on trending topics which have been upvoted by its users. It might be of extreme help in sociology research, due to its ability to display what leads to the concerns of the public; the chatbot is limited to performing the search and is unable to build up a conversation as ChatGPT does.

Microsoft Copilot is AI Chatbot for Windows. It is functioning in connection with other Microsoft products; the latter somewhat limits its application by vast majority of users.

You.com is a conversation-type Chatbot. It gives sources for its answers and suggests the links to them. If a quick answer is required, the Chat can provide it in a brief form.

Character AI is a chatbot for character-based conversations. The Chat assists users in choosing their character based on the unique features of their personality, save their memories, interests and way of talking. It enables talking to other characters.

Among 10 popular Chatbots listed above, ChatGPT tends to be the most fitting in terms of addressing the demands of the educators' professional tasks as it enables free access, includes the broadest number of functions, is not restricted to an operation system, and supplies information from a vast number of sources. So, ChatGPT can be recommended as the most appropriate for usage in education regardless of the country the educator is working from, and operation system available to them.

AI tools enter various spheres of human activity. Extensive applications of AI-aided tools have been reported in the healthcare sector, finance, legislation, politics, and education. Digital self-help programmes without human guidance [3] were initiated by online consulting in psychiatry; the initiative turned out to be a success, and in 2023, according to the experiment on mental health, the users of the support platform that applied ChatGPT-3, rated AI-generated suggestions to their mental health concerns higher than support provided by a human solely [4]. Also, recent studies are reporting the options to apply ChatGPT in secondary and tertiary levels

of health care [5]. Additionally, ChatGPT, based on GPT-4, earned an almost ideal 99% score concerning financial literacy posing an emergent ability of an advisor in financial matters [6]. AI-mediated communication in the legislation domain positively impacts political representation as well as improves the transparency of legislators' activities and thus can contribute to building up the trust for election candidates [7]. A wide scope of ChatGPT application points to its quite successful performance in terms of how people can obtain some medical advice, get access to the legislation and its interpretation, learn the basics of financial competence etc. So, AI tool – ChatGPT is currently aiding health care services, the finance sector as well as the legislation domain, and education is not an exception.

The potential of AI in education through the application of AI-based tools such as intelligent tutoring [8] and automated grading systems had been described before ChatGPT's public domain availability. In 2018 M. Montebello presented AI as a valuable tool to improve the already existing instruments for e-learning and technology-enhanced learning in higher education and their contribution to the life-long education of professionals. The research 'AI Injected e-learning' addressed a number of issues, among which there were methods like crowdsourcing via social networks, user profiling through machine learning techniques, and personal learning portfolios [9]. AI has been described as a tool with considerable potential of improving Learning Management Systems (LMS) which are extensively employed in distance education processes [10]. The Ukrainian educational environment has also been marked by AI efficacy research where AI is described as an effective media to assist lecturers in generating teaching material for conducting lectures [11]. ChatGPT, as an AI tool, was also the subject of the research outlining the thoughts of educators as well as students on the potential application of the tool in teaching English for engineering classes [12].

Despite the growing popularity of ChatGPT in health care, legislation, and finance, the tendencies in education split into two directions. On the one hand, there are studies covering concerns of academia about ChatGPT or other AI tools as a direct threat to academic integrity, negatively impacting the development of creative and critical thinking [13], [14]. The researchers claim ChatGPT to be an issue considerably hindering the creative potential of non-native English learners while producing texts especially of those with weaker writing skills [15]. So as a result, some educational establishments even ban its usage [14].

However, there is another direction embracing the researchers who point out not only its potential in improving education, as it has been recommended with e-learning and intelligent tutoring, but also the revolutionising contribution of ChatGPT usage, viewing it as a tool to bring multiple advantages, namely: assisting teachers and students in their teaching and learning practices, such as preparing teaching materials, creating quizzes, etc. [16]. Undoubtedly, the idea of implementing ChatGPT or any other AI tool requires alterations in teaching philosophy, when educators are "upskilling their competencies and practices to meet the new demands of technology", e.g. they can integrate a strategy of using debates together with written assignments, the performance of which might be assisted by AI-generated tools [17]. Additionally, the team of scientists offers the stakeholders, such as higher education institutions, to implement the sensitisation programs to develop the proper guidelines for AI chatbot usage constructing manuals to assist the teachers with ChatGPT [16].

A beneficial aspect of AI application in academia is challenged by J. L. Steel who argues that ChatGPT and other generative AI tools are undoubtedly threatening education of the present time in terms of measurement, data precision, and questioning the efficacy of teaching as well as learning skills [18]. Moreover, researchers investigating ChatGPT's ability to generate human understanding in terms of beauty, humour, surprise, irony, envy, etc. reported the failure of the latter to do so, due to its efficacy in generating human language patterns rather than performing cognitive functions responsible for shaping human understanding [19], [20], [21]. This implies that adhering to entirely AI-based decisions, like ChatGPT-assisted

assessment of creative works poses a certain threat to unbiased assessment because of its compromised ability to recognise the above-mentioned human cognitive functions. In other words, being aware of the fact that ChatGPT is incapable of human thinking, research revealed that educators, of e.g. Master's Programme, might use the method of comparing AI-generated solutions to the problems stated, to those offered by the students [18].

It is noteworthy that apart from the studies welcoming ChatGPT to the education sector, some scientists opine that regardless of the crucial role of teachers [22], [23] as the final instance in verifying the task, its generation as well as its completion, they may apply AI tools in designing preliminary teaching activities connected with the collection of the study material, sorting it out etc. So there is still little information on how to use AI tools as teachers' assistants as well as there is limited knowledge about employment of AI-based tools in other sectors of education [24].

The study addresses the mapping of language educators' acquired competence and their intention to familiarize themselves with ChatGPT for the purposes of task design and assessment. The **purpose** of this paper, therefore, is to analyse insights and feedback provided by EFL and ESP instructors on the use of ChatGPT in English language teaching and assessment and to underscore challenges and concerns associated with its application.

2. RESEARCH METHODS

Mixed methods research was performed in this study, which combines elements of both quantitative and qualitative research to gain a more complete picture of the English language instructors' feedback obtained through a Google Form survey. The survey on the use of ChatGPT for assessment in English language teaching was conducted from January to March 2024. A total of 36 respondents participated in the survey, most of whom are EFL and/or ESP professors of higher education institutions.

The survey aimed to achieve the following objectives:

- 1) Gather information about the affiliation of participants to understand the diversity of backgrounds and contexts in which ChatGPT is being used for assessment in English language teaching;
- 2) Determine the level of teaching experience of participants to gauge the range of perspectives and expertise in utilizing technology like ChatGPT for educational purposes;
- 3) Assess participants' familiarity with ChatGPT to understand the extent of their exposure and experience with this technology;
- 4) Determine whether participants have integrated ChatGPT into their English language teaching activities and gather insights into the specific methods and approaches employed;
- 5) Identify the language skills assessed with the assistance of ChatGPT, including grammar, vocabulary, reading comprehension, writing, listening comprehension, and speaking;
- 6) Assess the perceived accuracy of ChatGPT-generated tasks for evaluating students' language skills, providing insights into its reliability as an assessment tool.
- 7) Identify any challenges or concerns encountered by participants while using ChatGPT for assessment, facilitating the identification of areas for improvement and support;
- 8) Gauge participants' overall satisfaction with their experience of using ChatGPT for assessment on a numerical scale, allowing for quantitative analysis of user satisfaction;
- 9) Determine whether participants require guidelines and recommendations for creating tasks with the assistance of ChatGPT, identifying potential areas where additional support may be beneficial;
- 10) Invite participants to provide any additional comments or insights regarding their experiences with using ChatGPT for assessment, enriching the understanding of its practical implementation and impact in English language teaching.

3. FINDINGS

This survey is part of a research project investigating various aspects of integrating ChatGPT into English language teaching practices. By collecting responses from educators who have employed ChatGPT in their teaching activities, we aimed to understand its impact, effectiveness, and potential challenges. This survey will inform the development of guidelines and recommendations for integrating ChatGPT into language assessment practices, ultimately enhancing its utility and relevance in educational settings. Thus, exploring the potential and challenges of using this advanced AI tool to enhance language assessment, the study provided valuable insights into how educators perceive and utilize AI in their teaching practices. Understanding these perspectives may help in designing better AI tools that align with pedagogical goals and address the specific needs and concerns of teachers.

According to the set objectives, the following findings have been revealed.

1. Most respondents are affiliated with various Ukrainian higher education institutions, including Taras Shevchenko National University of Kyiv, Yaroslav Mudryi National Law University, Bohomolets National Medical University, Borys Hrinchenko Kyiv University, etc. There are respondents affiliated with international universities such as the University of Richmond School of Law, Georgetown University Law Center, Georgetown Law, Suffolk University, Georgetown Law School, Nicolaus Copernicus University in Toruń, Poland, etc. Some respondents indicated that they are self-employed, which could include independent legal English teachers or individuals working in fields not directly associated with educational institutions.

Overall, the responses demonstrate a diverse range of affiliations, spanning various educational institutions, departments, and professional roles. This diversity allows us to suggest that the survey respondents represent a broad spectrum of backgrounds and experiences within the field of English language teaching and legal education.

2. As for the years of teaching experience, the survey shows the following results (Fig. 1).

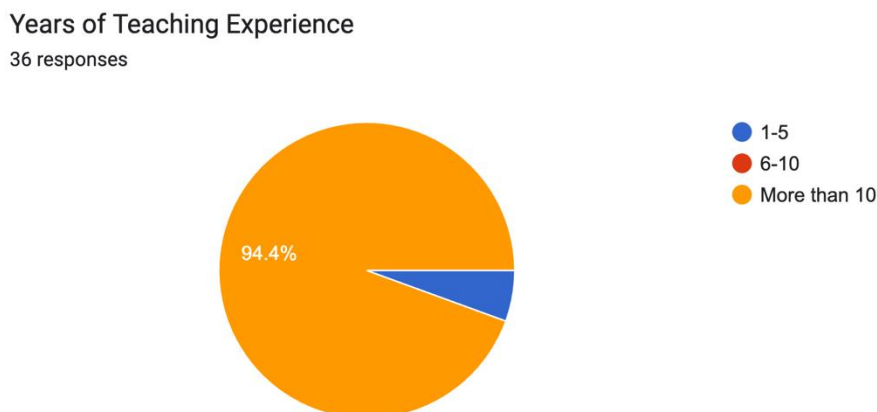


Figure 1. Responses to the question on years of teaching experience

The vast majority of respondents, accounting for 94.4% of the total, has more than 10 years of teaching experience. The high percentage suggests that a significant proportion of survey participants are experienced educators who have been teaching for over a decade. We presume that their wealth of experience can provide valuable insights into the integration of ChatGPT in English language teaching and assessment. Respondents with 1-5 years of teaching experience represent 5.6% of the total. While this group is less prevalent, their perspectives are

still valuable, particularly in understanding the experiences and viewpoints of educators who are newer to the profession.

3. The analysis of the responses to the question on “How familiar are you with ChatGPT?” reveals varying levels of familiarity among the survey participants (Fig. 2):

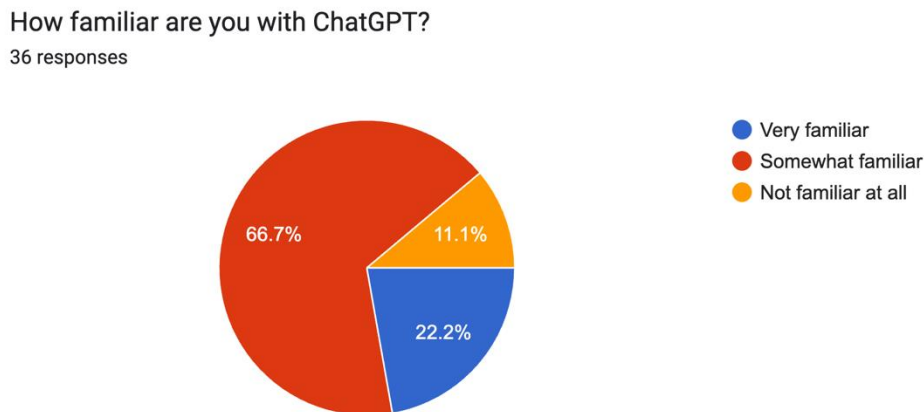


Figure 2. Responses to the question on the levels of familiarity with ChatGPT

The majority of respondents, accounting for 60% of the total, indicated that they are somewhat familiar with ChatGPT. Thus, we can suggest that a significant portion of the participants have some knowledge or experience with ChatGPT but may not be extensively versed in its capabilities or applications. A smaller proportion of respondents, comprising 20% of the total, stated that they are very familiar with ChatGPT. This group likely consists of individuals who have a deeper understanding of ChatGPT, its functionalities, and its potential applications to language teaching and assessment. A minority of respondents, representing 11% of the total, indicated that they are not familiar at all with ChatGPT. This category includes participants who have little to no prior exposure or knowledge of ChatGPT and its use in educational contexts.

Overall, the responses to this question demonstrate a range of familiarity levels with ChatGPT among the survey participants. While the majority have at least certain level of familiarity, there is also a notable portion who may require more information or guidance to understand the technology and its implications for language teaching and assessment. This diversity in familiarity levels underscores the importance of providing adequate context and support when discussing ChatGPT in the context of English language teaching.

4. The analysis of the responses to the next three questions (“Have you integrated ChatGPT into your English language teaching activities?”, “If you have integrated ChatGPT into your English language teaching activities, please describe how you have used ChatGPT in your teaching.”, and “If you have NOT integrated ChatGPT into your English language teaching activities, please provide the reason why not”) provides insights into the adoption of ChatGPT among the survey participants.

The majority of respondents, comprising 55.6% of the total, indicated that they have not integrated ChatGPT into their English language teaching activities. The responses to the question about why respondents have not integrated ChatGPT into their English language teaching activities provide valuable insights into the perceived barriers and considerations surrounding the adoption of this technology. Thus, these responses highlight a range of factors influencing educators’ decisions not to integrate ChatGPT into their teaching activities, including pedagogical beliefs, concerns about credibility and reliability, lack of knowledge or

resources, and preferences for traditional teaching methods. We can suggest that addressing these concerns and providing guidance and support for effectively incorporating AI technologies into language teaching could help encourage broader adoption in the future.

A smaller proportion of respondents, accounting for 44.4% of the total, stated that they have integrated ChatGPT into their English language teaching activities. The survey reveals that ChatGPT is used for the following activities: generating and evaluating answers, creating language exercises, language correction and analysis, writing companion, lesson planning and content creation, assessment and test creation, ESP course development, etc. Overall, the responses highlight the versatility of ChatGPT in English language teaching, ranging from language practice and correction to content creation, assessment, and lesson planning. Educators leverage ChatGPT to enhance various aspects of language instruction and student learning, demonstrating its potential as a valuable tool in the classroom.

5. The responses to the question about what educators have assessed with the help of ChatGPT provide insights into the specific language skills and areas of focus addressed through the integration of this technology into ESL classes. Figure 3 illustrates that ChatGPT is used to assess grammar (12 responses), vocabulary (13 responses), reading comprehension (6 responses), listening comprehension (2 responses), writing (11 responses), and speaking (4 responses).

What have you assessed with the help of ChatGPT? (Select all that apply or skip the question if you have not applied ChatGPT to your ESL classes)

17 responses

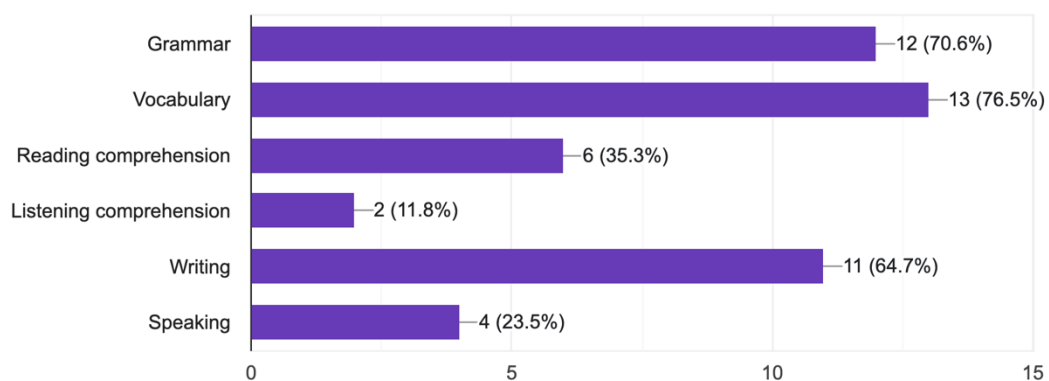


Figure 3. Responses to the question on what educators have assessed with the help of ChatGPT

Overall, the responses demonstrate a diverse range of language skills and areas that educators have addressed through the integration of ChatGPT into ESL classes. Thus, we can suggest that ChatGPT offers versatile capabilities for assessing and enhancing various aspects of English language proficiency.

6. The responses to the next question provide insights into educators' perceptions of the accuracy of ChatGPT-generated tasks for assessing students' language skills (Fig. 4).

How accurate do you find ChatGPT-generated tasks for assessing students' language skills?

36 responses

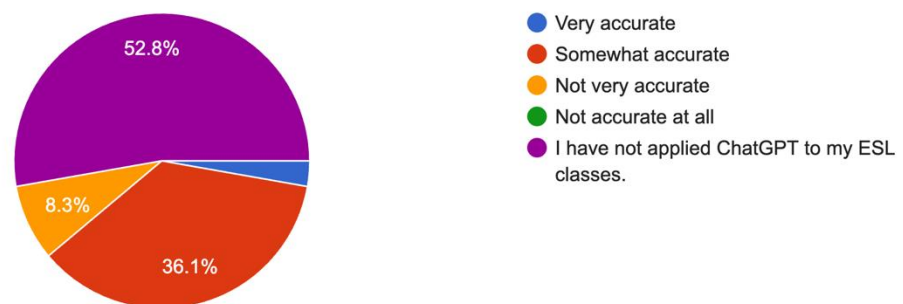


Figure 4. Responses to the question on the accuracy of ChatGPT-generated tasks for assessing students' language skills

Nineteen respondents have not applied ChatGPT to their ESL classes. These individuals have not had direct experience with using ChatGPT for assessing students' language skills and thus cannot provide feedback on its accuracy. Thirteen respondents perceive ChatGPT-generated tasks to be somewhat accurate for assessing students' language skills. Three respondents indicate that ChatGPT-generated tasks are not very accurate for assessing students' language skills. Only one respondent finds ChatGPT-generated tasks to be very accurate for assessing students' language skills.

Overall, while some respondents perceive ChatGPT-generated tasks to be accurate or somewhat accurate for assessing language skills, others express reservations about its accuracy. The diverse range of responses highlights the need for further exploration and evaluation of ChatGPT's effectiveness in assessing language proficiency.

7. The responses to the next question provide insights into the challenges or concerns encountered by educators when using ChatGPT for assessment purposes (Fig. 5).

Have you encountered any challenges or concerns while using ChatGPT for assessment?

36 responses

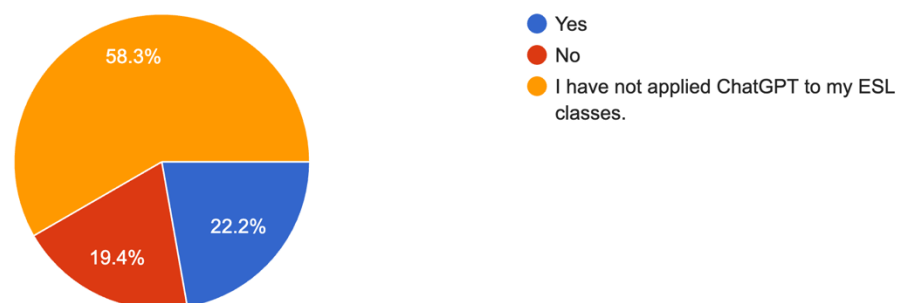


Figure 5. Responses to the question on the challenges or concerns encountered by educators when using ChatGPT for assessment purposes

Eight respondents indicate that they have encountered challenges or concerns while using ChatGPT for assessment. Seven respondents report that they have not encountered any challenges or concerns while using ChatGPT for assessment. Twenty-one respondents have not applied ChatGPT to their ESL classes.

The responses to the question about challenges or concerns encountered while using ChatGPT for assessment highlight several key issues. Respondents note challenges in creating tasks that effectively check subtle differences in meanings or achieve desired results, indicating limitations in ChatGPT's ability to meet specific instructional needs. Several respondents mention the need for thorough proofreading of tasks generated by ChatGPT and the importance of specifying context, particularly in English for Specific Purposes settings, to ensure relevance and accuracy. Some respondents mention instances where ChatGPT fails to provide accurate or relevant information, lacks background information, transitions, details, or explanation in writing tasks, indicating shortcomings in content generation. One respondent notes that ChatGPT tends to formulate comprehension questions in a straightforward manner by simply paraphrasing information from the text, lacking intricacy or depth. Another respondent expresses concerns about the reliability of ChatGPT and larger ethical issues, particularly regarding privacy concerns associated with requiring students to create accounts and use ChatGPT. Others cite a lack of expertise in working with ChatGPT and the need for guidelines for its application, underscoring the importance of professional development and support in effectively integrating AI technologies into teaching practice.

Overall, the responses highlight various challenges and concerns related to the use of ChatGPT for assessment purposes, including limitations in task complexity, reliability, accuracy, and adaptability to diverse instructional contexts. In this regard, we can assume that addressing challenges and concerns, as well as sharing best practices and guidelines, can help educators maximize the benefits of integrating ChatGPT into their assessment practices effectively.

8. The analysis of the responses on educators' overall satisfaction with their experience of using ChatGPT for assessment in English language teaching reveals the following insights (Fig. 6).

On a scale from 1 to 10, how satisfied are you with your overall experience of using ChatGPT for assessment in English language teaching? (1 - ext...you have not applied ChatGPT to your ESL classes
18 responses

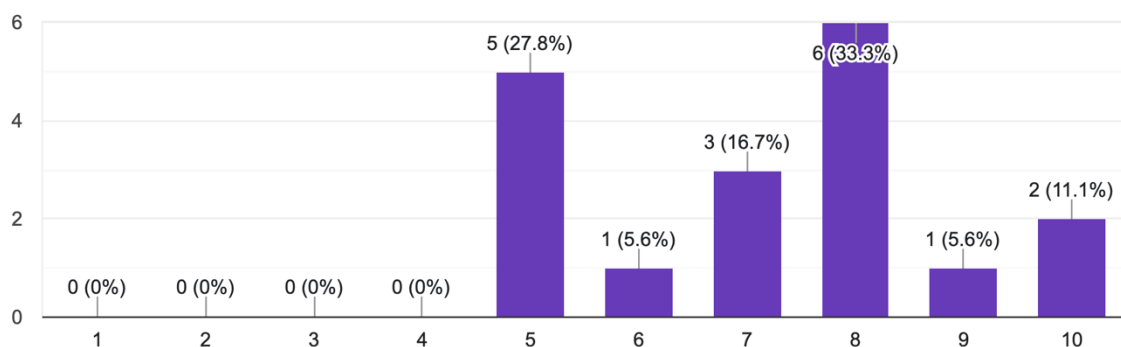


Figure 6. Responses to the question on educators' overall satisfaction with using ChatGPT for assessment in English language teaching

Eighteen respondents did not provide a response to the question, indicating that they have not applied ChatGPT to their ESL classes for assessment purposes. Therefore, they cannot provide feedback on their satisfaction with using ChatGPT in this context. Six respondents rated their overall satisfaction with using ChatGPT for assessment relatively high, with ratings of 8, 9, or 10. This suggests that these educators have had positive experiences with ChatGPT and

are generally satisfied with its effectiveness in assessing English language skills. Five respondents rated their overall satisfaction with using ChatGPT for assessment at a moderate level, with ratings of 5, 6, or 7. This indicates that while these educators may have found some value in using ChatGPT, there may also be areas for improvement or challenges that have affected their satisfaction.

Overall, the responses reflect a range of satisfaction levels among educators who have applied ChatGPT to their EFL classes for assessment purposes. While some respondents express high satisfaction with their experience, others indicate more moderate levels of satisfaction. In this regard, we can presume that further exploration of the factors influencing satisfaction, such as task design, ease of use, and alignment with instructional objectives, could provide valuable insights for enhancing the effectiveness of ChatGPT in English language teaching.

9. The responses to the question regarding the perceived helpfulness of guidelines or recommendations on creating ChatGPT-generated test tasks for EFL/ESP courses, indicate a strong interest among educators in receiving guidance in this area (Fig. 7).

Would guidelines or recommendations on how to create Chat GPT-generated test tasks for EFL /ESP courses be helpful for you?

36 responses

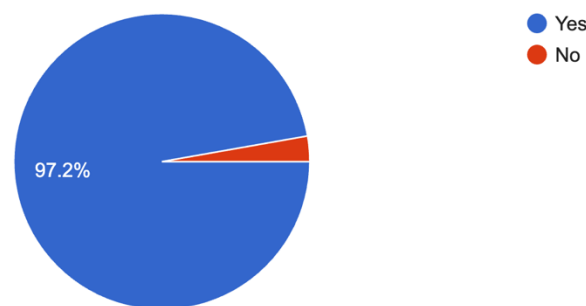


Figure 7. Responses to the question on the educators' need for guidelines or recommendations on creating ChatGPT-generated test tasks for EFL/ESP courses

Thirty-five respondents expressed that guidelines or recommendations on creating ChatGPT-generated test tasks for EFL/ESP courses would be helpful. This overwhelming majority suggests that educators recognize the potential benefits of having structured guidance or best practices to follow when incorporating ChatGPT into their assessment practices. The desire for guidelines may stem from a need for clarity on task design, alignment with curriculum objectives, and effective implementation strategies. Only one respondent indicated that they do not believe guidelines or recommendations would be helpful. While this response represents a minority viewpoint, it still suggests that some educators may feel confident in their ability to create ChatGPT-generated test tasks independently or may have reservations about the utility of guidelines in this context.

10. The responses to the last question of the survey provide additional insights and comments regarding educators' experiences with using ChatGPT for assessment. Thus, many educators express an interest in learning more about ChatGPT's potential functions and incorporating ChatGPT-generated tests into their classes. There is a clear need for guidelines and recommendations from experienced professionals who have used ChatGPT for writing and testing purposes. Educators believe that guidelines based on research would be beneficial in effectively utilizing ChatGPT in their teaching practices. They also suggest avenues for further

learning, such as webinars, to gain knowledge and skills in using ChatGPT for assessment. This indicates a willingness to engage in professional development activities to enhance their proficiency in integrating AI technologies into their teaching. Some educators acknowledge the potential benefits of using ChatGPT for assessment, such as generating exercises and saving time in content development. However, there is variability in the frequency of use, with some not using it regularly. Concerns are also raised regarding the accuracy of exercises generated by ChatGPT, particularly for training vocabulary. Educators note the need for editing due to inaccuracies or incorrect explanations generated by AI.

Overall, the responses highlight educators' interest in leveraging ChatGPT for assessment purposes, along with the need for support, guidance, and professional development opportunities to maximize its effectiveness in English language teaching.

Thus, the research offers a detailed analysis of the accuracy and effectiveness of ChatGPT-generated assessment tasks across various language skills, including grammar, vocabulary, reading comprehension, writing, listening comprehension, and speaking. This thorough evaluation contributes new knowledge about the capabilities and limitations of ChatGPT in language education.

Focusing specifically on EFL and ESP courses, the study adds a layer of specificity and relevance that enhances the understanding of AI's role in different educational settings. By addressing ethical concerns, privacy issues, and the need for human oversight in AI-generated tasks, the research brings a balanced perspective to the discussion, which is critical for the responsible integration of AI in education.

The study also identifies a strong demand among educators for guidelines and professional development related to using ChatGPT. Highlighting this need and suggesting avenues for support and training represent a forward-thinking approach that anticipates and addresses practical implementation challenges. Furthermore, the involvement of participants from various institutions, including the University of Richmond School of Law, Georgetown University Law Center, and Suffolk University, adds a collaborative dimension to the research. This diversity of input enriches the findings and broadens their applicability.

4. CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

The research explores the innovative use of ChatGPT, an advanced language model, in language assessment. While AI has been used in education, applying ChatGPT specifically for creating and evaluating language assessment tasks is relatively new. The survey phase of the study aimed to gather first-hand experiences and feedback from educators who have integrated ChatGPT into their teaching, and who expressed concerns preventing them from applying it, providing unique insights into its real-world applicability, benefits, and challenges, which are less commonly explored in existing literature.

The participants of the survey are mostly accomplished instructors of English with more than ten-year teaching experience, who are either affiliated with a higher education institution in Ukraine, the EU or the USA or are self-employed. Their expertise in English language teaching and assessment may be the reason for seeking new approaches and innovative tools to apply in their job.

The number of the respondents who are familiar with ChatGPT only up to a point or are not familiar with ChatGPT significantly prevails those who feel confident with ChatGPT's functions and are fully aware of its potential. More than a half of the former admitted that they have not attempted to use this AI tool in their teaching practice due to a number of constraints the most noteworthy of which are lack of knowledge and guidelines for its effective implementation and doubts as to credibility of its output. Those educators who use ChatGPT, however, admit that it may be quite beneficial as an assistant providing support in the issues of

lesson planning, content creation, task design and assessment. Though the responses display a wide range of language skills which can be formed through ChatGPT's assistance, the areas where educators find ChatGPT the most supportive for teaching and assessment include vocabulary, grammar and writing. This underscores the tool's potential for designing productive skills development and assessment tasks.

The majority of respondents express doubts as to the accuracy of the tasks generated by ChatGPT for assessment purposes, which necessitates thorough examination of the constraints impeding task accuracy, and development of criteria for ChatGPT-generated tests evaluation. Other challenges English language instructors have been exposed to include the need for careful proofreading of the output by a human teacher, inaccurate and biased information in generated content, lack of task complexity. Nevertheless, the overall level of satisfaction with using ChatGPT for teaching and assessment among those educators who have this experience range from moderate to high, highlighting undisputable value this AI tool may have for foreign language teaching practice. At the same time, the vast majority of participants in the survey stressed the need for guidelines and recommendations on creating ChatGPT-generated test tasks for EFL/ESP courses. Taking into account the insights derived from the survey results analysis, further research will be undertaken to develop rigorous criteria for assessing the quality of the tasks generated by ChatGPT. In addition, elaboration of ethical policies regarding the application of AI to prepare assessment tasks could be prospective to address concerns and fears of educators towards integrating AI in academia as well as for popularization of its benefits for teaching and assessment.

Moreover, this research lays the groundwork for future studies and developments. It highlights current challenges and areas for improvement, guiding future innovations in AI-driven educational technologies. The findings of this research have implications for English language teaching worldwide, as English is a global lingua franca. Insights gained from this study can be applied in diverse educational contexts, such as providing personalized learning experiences and supporting diverse learning styles of students catering to their specific needs and learning paces, offering more adaptive and interactive tools and resources for native and non-native speakers, providing English language teachers with approaches for using AI technologies in a personalized way to address their unique challenges in planning lessons and designing tests.

Acknowledgments

We would like to express our sincere gratitude to our colleagues who actively participated in the survey. Your insights and contributions have been invaluable to this research. In particular, we extend our thanks to the participants from the University of Richmond School of Law, Georgetown University Law Center, Suffolk University, and many others for their enthusiastic involvement and support. Your collaboration has significantly enriched the findings of this study.

REFERENCES (TRANSLATED AND TRANSLITERATED)

- [1] A. Kukulska-Hulme, E. Beirne, G. Conole, E. Costello, T. Coughlan, R. Ferguson, E. FitzGerald, M. Gaved, Ch. Herodotou, W. Holmes, C. Mac Lochlainn, M. N. G. Mhichíl, B. Rienties, J. Sargent, E. Scanlon, M. Sharples, D. Whitelock. *"Innovating Pedagogy 2020: Open University Innovation Report"*, Institute of Education and Technology (The UK), National Institute for Digital Learning, Dublin City, Ireland, p. 10, 2020. [Online]. Available: https://www.researchgate.net/publication/338991922_Innovating_Pedagogy_2020_Open_University_Innovation_Report_8#fullTextFileContent

- [2] S. Sinha, L. Burd, J. Du. Preez, “How ChatGPT could revolutionize academia”, *IEEE Spectrum*, Feb. 2023. [Online]. Available: <https://spectrum.ieee.org/how-chatgpt-could-revolutionize-academia>
- [3] G. Andersson, N. Titov, B. F. Dear, A. Rozental, P. Carlbring, “Internet-delivered psychological treatments: from innovation to implementation”, *World Psychiatry*, vol. 18 (1), 2019, pp. 20–28. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/wps.20610>
- [4] B. Biron, “Online mental health company uses ChatGPT to help respond to users in Experiment – raising ethical concerns around healthcare and AI technology”, *Business Insider*, 2023. [Online]. Available: <https://www.businessinsider.com/company-using-chatgpt-mental-health-support-ethical-issues-2023-1>.
- [5] A. Sarraju, D. Bruemmer, E. Van Iterson, L. Cho, F. Rodriguez, L. Laffin, “Appropriateness of Cardiovascular Disease Prevention Recommendations Obtained From a Popular Online Chat-Based Artificial Intelligence Model”, *JAMA*, vol. 329, pp. 842–844, 2023. doi:10.1001/jama.2023.1044.
- [6] P. Niszczota, S. Abbas, “GPT has become financially literate: Insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice”, *Finance Research Letters*, vol. 58 (Part A), no. 104333, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S1544612323007055?via%3Dihub>
- [7] S. Kreps, M. Jakesch, “Can AI Communication Tools Increase Legislative Responsiveness and Trust in Democratic Institutions?”, *Government Information Quarterly*, vol. 40(3), no 101829, June 2023, [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0740624X23000291#preview-section-cited-by>
- [8] S. Steenbergen-Hu, H. Cooper, “A Meta-analysis of the Effectiveness of Intelligent Tutoring Systems on College Students’ Academic Learning”, *Journal of Educational Psychology*, vol.106, pp. 331–347, 2014.
- [9] M. Montebello, *AI injected e-learning: The future of online education*, Springer, 2018, doi:10.1007/978-3-319-67928-0.
- [10] D.-C. Arghir, “Implementation of Learning Management Systems with Generative Artificial Intelligence Functions in the Post-pandemic Environment”, *Information Technologies and Learning Tools*, vol. 100, no 2, 2024 [Online]. Available: <https://journal.iitta.gov.ua/index.php/itlt/issue/view/129>
- [11] O. Spivakovsky, S. Omelchuk, V. Kobets, Nataliia V. Valko, Daria S. Malchykova, “Institutional Policies on Artificial Intelligence in University Learning, Teaching and Research”, *Information Technologies and Learning Tools*, vol 97, no 5, 2023. [Online]. Available: <https://journal.iitta.gov.ua/index.php/itlt/issue/view/126>
- [12] O. Synekop, I. Lytovchenko, Y. Lavrysh, V. Lukianenko, “Use of Chat GPT in English for Engineering Classes: Are Students’ and Teachers’ Views on Its Opportunities and Challenges Similar?” *International Journal of Interactive Mobile Technologies (iJIM)*, 18(3), pp. 129–146, (2024). <https://doi.org/10.3991/ijim.v18i03.45025>.
- [13] A. Shiri, “ChatGPT and Academic Integrity”, *Information Matters*, vol. 3, no 2, Feb. 2023. [Online]. Available: <https://ssrn.com/abstract=4360052>
- [14] K. Huang, “Alarmed by AI Chatbots, Universities Start Revamping How They Teach”, *New York Times*, Jan. 16, 2023. [Online]. Available: <https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html>
- [15] W. Liang, M. Yuksekgonul, Y. Mao, E. Wu, J. Zou, “GPT detectors are biased against non-native English writers”, *ArXiv*, vol. 2, Jul. 2023. [Online]. Available: <https://arxiv.org/pdf/2304.02819.pdf>.
- [16] A. Herft, *A Teacher's Prompt Guide to ChatGPT: Aligned with 'What Works Best'*. CESE NSW "What Works Best in Practice," 2023. [Online]. Available: https://drive.google.com/file/d/15qAxnUzOwAPwHzoaKBJd8FAgiOZYcIqx/view?fbclid=IwAR2fRdL5ggq4zU-81FiI8j4BAOp5HqWHC_Ecy2sqKk4EiWXL0FKa5GVz5dE
- [17] A. Tlili, B. Shehata, M.A. Adarkwah, A. Bozkurt, D.T. Hickey, R. Huang, B. Agyemang, “What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education”, *Smart Learning Environments*, vol. 10(1), no 15, 2023. [Online]. Available:https://link.springer.com/article/10.1186/s40561-023-00237-x?utm_source=getftr&utm_medium=getftr&utm_campaign=getftr_pilot
- [18] J.L. Steele, “To GPT or not GPT? Empowering our Students to Learn with AI”, *Computers and Education: Artificial Intelligence*, vol. 5, no. 100160 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666920X23000395>
- [19] I. Bogost, “ChatGPT is dumber than you think”, *The Atlantic*, Dec. 2022. [Online]. Available: <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/>
- [20] C. Newport, “What kind of mind does ChatGPT have?”, *The New Yorker*, April 2023. [Online]. Available: <https://www.newyorker.com/science/annals-of-artificial-intelligence/what-kind-of-mind-does-chatgpt-have>

- [21] M. Schneider, “Do we need humans anymore?”, U.S. Department of Education, Institute of Education Sciences. Jan. 2023, [Online]. Available: <https://ies.ed.gov/director/remarks/01-04-2023.asp>
- [22] R. Luckin, K. George, and M. Cukurova, “AI for school teachers”, London, UK: CRC Press, 2022.
- [23] S. Seufert, J. Guggemos, M. Sailer, “Technology-related knowledge, skills, and attitudes of pre-and in-service teachers: The current situation and emerging trends”, *Computers in Human Behavior*, vol. 115, no. 106552, Feb. 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0747563220303022>
- [24] S. Kim, Y. Jang, S. Choi, W. Kim, H. Jung, S. Kim, H. Kim, “Analyzing teacher competency with TPACK for K-12 AI education” *KI-Künstliche Intelligenz*, vol. 35 (2), pp. 139–151, June 2021. [Online]. Available: <https://pure.korea.ac.kr/en/publications/analyzing-teacher-competency-with-tpack-for-k-12-ai-education>

Text of the article was accepted by Editorial Team 11.05.2024

CHATGPT У НАВЧАННІ ІНОЗЕМНИХ МОВ ТА ОЦІНЮВАННІ: ДОСВІД ВИКЛАДАЧІВ АНГЛІЙСЬКОЇ МОВИ

Інна Зяярна

кандидат педагогічних наук, доцент,
доцент кафедри іноземних мов
Навчально-наукового інституту права
Київський національний університет імені Тараса Шевченка, м. Київ, Україна;
провідний науковий співробітник відділу компаративістики інформаційно-освітніх інновацій,
Інститут цифровізації освіти НАПН України, м. Київ, Україна
ORCID ID 0000-0002-9464-096X
izaiarna@gmail.com

Олена Жигadlo

кандидат філологічних наук, доцент,
доцент кафедри іноземних мов
Навчально-наукового інституту права
Київський національний університет імені Тараса Шевченка, м. Київ, Україна
ORCID ID 0000-0002-1605-7242
olena.zhygado@gmail.com

Ольга Дунаєвська

кандидат філологічних наук, доцент,
доцент кафедри іноземних мов
Навчально-наукового інституту права
Київський національний університет імені Тараса Шевченка, м. Київ, Україна
ORCID ID 0000-0002-0915-428X
o.dunayevska@gmail.com

Анотація. У статті представлено результати першого етапу дослідження особливостей застосування ChatGPT у розробці інструментів оцінювання для курсів англійської мови як іноземної. У статті проаналізовано досвід викладачів англійської мови щодо використання ChatGPT у навчальній та оцінювальній практиці. Результати отримані за допомогою опитування, з метою встановлення ступеня обізнаності викладачів з можливостями використання ChatGPT для викладання іноземної мови та механізмами розробки тестових завдань за допомогою ШІ. Більшість учасників - досвідчені викладачі англійської мови з вищих навчальних закладів України, ЄС та США. Опитування продемонструвало різні рівні опанування респондентами функцій ChatGPT. Дослідження виявило, що значна кількість викладачів або не знайомі з ChatGPT, або мають обмежені знання про його можливості. Причини відмови використовувати ChatGPT у викладанні насамперед пов'язані з відсутністю рекомендацій щодо його ефективного використання, а також питаннями академічної доброчесності. Однак викладачі, які користуються ChatGPT, визнають його потенціал як помічника в плануванні уроків, формуванні змісту навчального процесу, розробці завдань та під час оцінювання лексичних і граматичних навичок та писемного мовлення. Незважаючи на переваги цього інструменту, респонденти мають застереження щодо точності завдань,

створених ChatGPT для оцінювання. Серед виявлених викликів - необхідність ретельного редагування, недостовірна інформація та недостатня складність завдань. Проте загальний рівень задоволення серед викладачів, які мають досвід впровадження ChatGPT, варіює від помірного до високого, що підтверджує його цінність у процесі викладання іноземних мов. Тож на етапі опитування було виявлено зростаючий інтерес як серед українських, так і серед іноземних викладачів до потенціалу ChatGPT у викладанні іноземної мови, зокрема під час оцінювання, і підтверджено необхідність рекомендацій для створення ефективних тестових завдань, створених за допомогою ChatGPT.

У подальшому дослідженні планується розробити критерії оцінки якості завдань, згенерованих ChatGPT, створити рекомендації щодо формулювання запитів та запропонувати етичні принципи застосування ШІ для оцінювання, які б задовільнили потреби викладачів та сприяли інтеграції ШІ в академічне середовище.

Ключові слова: штучний інтелект (ШІ); ChatGPT; тестування та оцінка; іншомовна компетентність; опитування.

