2024 IEEE International Conference on Advanced Learning Technologies (ICALT) | 979-8-3503-6205-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICALT61570.2024.00062

# Enhancing Academic Performance with Generative AI-Based Quiz Platform

Chia-Kai Chang*
*Center for General Education*
*National Central University*
Taoyuan City, Taiwan
ckchang@ncu.edu.tw

Lee-Chia-Tung Chien
*International Graduate Program in Artificial Intelligence*
*National Central University*
Taoyuan City, Taiwan
a90799@gmail.com

***Abstract*—In this work, we build a QuizGPT web interface for learners to learn Python via quizzing and chatting. Additionally, we evaluate the impact of generative AI on learning. To effectively evaluate learners' learning achievements, the quizzes must be diverse, and the options should be challenging to ensure a complete understanding of the material. It is a time-consuming challenge for teachers to create adaptive quizzes based on learners' levels for individual learning needs. This study generates 366 quiz questions automatically via generative AI in three difficulty levels (easy, medium, hard), based on Python knowledge points. A positive correlation observed between the QuizGPT platform activities and learners' test scores. These activities include the number of quizzes attended, the correctness rates, and the frequency of participation in generative AI sessions. Additionally, QuizGPT platform activities could predict learners' test scores with a Root Mean Squared Error percentage (%RMSE) of 3.58%. Remarkably, interaction with generative AI proves to enhance Python programming skills more effectively than the number of quiz attempts. Survey analysis reveals that QuizGPT can reduce learning anxiety and enhance learning interest. Our findings suggest that generative AI, as implemented in the QuizGPT platform, is a potent tool for academic improvement, significantly enhancing self-regulation and engagement in learning Python programming.**

***Keywords*—Generative AI, Artificial intelligence generated content, ChatGPT, Quiz platform.**

## I. Introduction

ChatGPT, an acronym for Chat Generative Pre-Trained Transformer, is an artificial intelligence tool launched by OpenAI on November 20, 2022, as a web-based chatbot application[1], [2]. Chatbots are software applications based on artificial intelligence, engineered to simulate human conversation using text or audio. They deliver responses in natural language to human inputs, maintaining a conversational style. Generative AI (GAI) could create text, images, or sound. This technology may have a significant impact on education. The use of GAI for personalized education and enhancing learners' learning effectiveness have become areas of interest in digital education[3].

Practicing exercises during a course helps learners revisit and strengthen their memory after finishing a chapter. They are designed to complement the instructional materials. The exercises should include various types of questions, such as positive and negative statements, and the multiple-choice options should be appealing to improve the reliability and validity of the exercises. Adaptive exercises are necessary to meet the needs of learners at different levels and enable adaptive teaching. Creating multiple-choice questions manually is time-consuming, and it is difficult to offer personalized exercises efficiently for learners. This study focuses on the general education course "Python Programming" at a certain national university in Taiwan. The test format consists of single-answer multiple-choice questions, typically featuring a short segment of Python code. Learners are required to perform "thought programming," mentally following the program's logic to select the correct result from four options.

We created a learning platform called 'QuizGPT' for students to practice Python programming on the 'NCUEDU' educational experiment platform (https://ncuedu.tw/). The flowchart of quiz preparation, learner interaction in QuizGPT, and the machine learning analysis are shown in Fig. 1. In QuizGPT, if a student answers incorrectly, they are encouraged to try again. If the answer is correct, ChatGPT provides a detailed explanation of the correct answer. The quizzes aim to assist students in memorizing and comprehending key knowledge points. We encourage students to ask questions to ChatGPT and utilize the interactive chat for gaining knowledge.

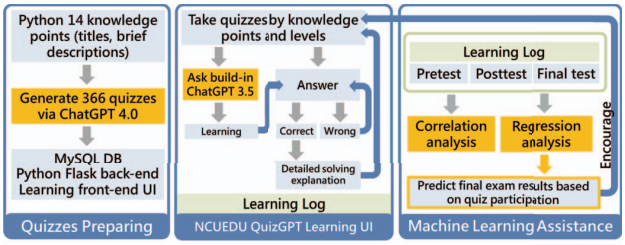



Fig. 1. The flowchart of quizzes generated, used, and data analyzation.

QuizGPT offers a comprehensive collection of Python Programming questions, automatically generated by GAI, spanning across 14 knowledge domains. With a total of 366 quizzes, they are categorized into easy, medium, and hard levels. This categorization enables learners to engage in adaptive learning tailored to their desired level of difficulty. As a learning support tool, QuizGPT facilitates seamless interaction between GAI and students, from question generation to real-time chatting. QuizGPT also facilitates the gathering of student learning behaviors for personalized analysis of quiz accuracy rates.

The research questions proposed in this study are summarized as follows:

RQ1: Is it feasible to use GAI-based Python multiple-choice quizzes to assess learners' learning outcomes?

2161-377X/24/$31.00 ©2024 IEEE
DOI 10.1109/ICALT61570.2024.00062

RQ2: Can learners' engagement in GAI-based quizzes improve learning outcomes?

RQ3: Does interaction with generative AI enhance learners' learning outcomes?

RQ4: How do GAI-based learning support tools affect learners' learning attitudes and learning emotions?

## II. Literature Reviews

Extensive research has been conducted on the application of ChatGPT in developing assessments and the generation of quizzes. Natural language processing is useful in many areas like informatics, marketing, education, political science, and interdisciplinary domains.

Educators' reception of ChatGPT has been mixed, with praise and controversy surrounding it. Some experts propose using AI-powered generators in education to improve teaching and learning rather than banning them[4]. Teachers can employ ChatGPT for lesson planning, student evaluation, and professional growth tasks. However, it's important to note that ChatGPT cannot access real-time internet information like conventional search engines, and its knowledge is confined to what it learned before September 2021, leading to occasional inaccuracies in factual information [5]. Many studies are currently exploring how ChatGPT can be used in different types of education. Lee use ChatGPT as virtual teaching assistants in medical education to enhance students' learning experiences [6]. Boscardin using ChatGPT to simulate real-world clinical scenarios and patient interactions, allowing students to learn and practice in a safe and controlled environment [7]. Lee designed a automatic question generation (AQG) system using large language models (LLMs) for English education [8]. Hajj focusing on the influence of ChatGPT in a PHP programming course, leverages user studies to elucidate the AI's role in enhancing learning outcomes [9].

## III. Method

To evaluate the impact of AIGC quizzes on Python programming learning, it is essential to automatically generate Python quizzes via GAI. Furthermore, an interface needs to be designed to facilitate learner participation in answering these quizzes.

### A. QuizGPT Platform Design

The 'NCUEDU' website and 'QuizGPT' utilize the Python Flask framework for its back-end, coupled with the MySQL Community Edition database. The web service is hosted on Vultr with regular performance virtual server (Intel 1 vCPU, 2 GB memory, 55 GB SSD) and uses the Ubuntu 20.04 LTS x64 operating system.

### B. Quizzes Generation

We identified 14 crucial Python concepts based on the Python syntax document. Commands are given to the ChatGPT 4.0 API to automatically create quiz questions with three difficulty levels (easy, medium, hard). This process created 366 Python multiple-choice quizzes.

## IV. Study Description

### A. Study Design

We implemented the QuizGPT platform into "Python Programming" course offered by the General Education Center at a research-oriented university in northern Taiwan. A total of 46 participants with various academic backgrounds (from 15 departments) successfully participated in this study (33 males, 13 females; aged 20.3±0.9 years; age range: 19 to 22 years). We plan to validate the effectiveness of GAI-based quiz questions by analyzing students' pretest, posttest, and final examination scores. Additionally, we assessed the QuizGPT platform's effectiveness through self-administered questionnaires. By analyzing students' responses to this questionnaire, we evaluate how the GAI-based QuizGPT learning system impacts the enhancement of learning emotions, reduction of learning anxiety, and improvement of learning outcomes.

### B. Study Procedures

All participants were briefed about the experimental details and asked to complete an informed consent form in the first week. They were administered a pretest in the first week and took a posttest, semester examination, and a questionnaire regarding their experience with QuizGPT in the twelfth week.

## V. Result

We collected data on learners' quiz participation, quiz accuracy, ChatGPT interaction, pretest scores, posttest scores, learning performance, and learner satisfaction questionnaires regarding their experience with the QuizGPT platform. And use these data to explain the four research questions.

### A. Effectiveness of AIGC-based Quiz in Learning Assessment

We plotted a scatter plot of the accuracy of the quiz against the exam score shown in Fig. 2 (a) to address RQ1. With regression analysis, we observed a Pearson correlation of 0.42. This indicates that quizzes generated automatically by ChatGPT 4.0 possess similar validity to human-generated multiple-choice questions.

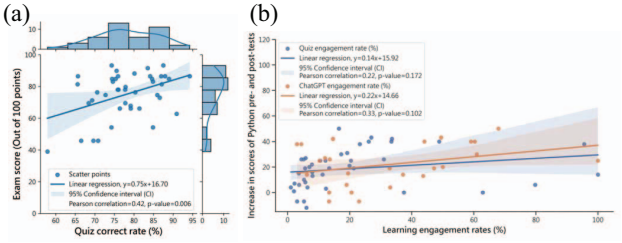



Fig. 2. (a) Scatter plot of quiz correct rate and exam score. (b) Scatter plot of quiz correct rate and exam score.

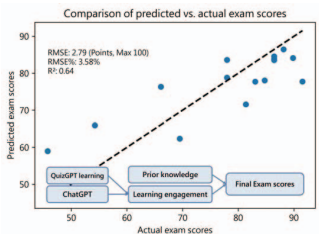
### B. Impact of QuizGPT participation on learning effectiveness

For RQ2 and RQ3, we investigated the impact of learners' participation in QuizGPT on their learning outcomes. We analyze the score growth between pre- and post-tests to account for the influence of prior knowledge on final grades. Thus, the research question evolves to investigate whether there is a correlation between students' involvement in QuizGPT learning activities and the increase in their scores from pre- to

post-testing. The QuizGPT learning activities are categorized into two types: the number of quiz questions answered and the number of interactions with ChatGPT. To compare these two activities on the same axis, we normalize each one and convert them into percentages on the x-axis. The y-axis represents the increase in scores between the pre-test and the post-test. The Fig. 2 (b) shows that learning engagement is divided into two categories: Quiz engagement (%) and ChatGPT engagement (%). Regardless of the type of learning activity, both Quiz engagement and ChatGPT engagement contribute to an increase in Python scores. It is worth mentioning that ChatGPT engagement seems to have a slightly greater influence on improving Python scores.

### C. Machine Learning Model for Grade Prediction

Using machine learning could predict exam scores based on learners' learning engagements (QuizGPT and ChatGPT). We collect data on learners' engagement in QuizGPT, pre and post-test scores, and learning exam scores. Creating a machine learning model using multiple linear regression. This study assume that exam scores are determined by two factors: prior knowledge of Python and learning engagements. Learning engagements are contributed by the number of quiz questions and ChatGPT interactions in QuizGPT. After multiple linear regression analyses, the final results are as Fig. 3.




**Machine learning linear regression model:**

$y = a_1x_1 + a_2x_2 + a_3x_3 + b$

$y$: Exam scores (Max 100)

$x_1$: Quiz participation counts ([56, 364])

$x_2$: ChatGPT participation counts ([0, 21])

$x_3$: Pretest scores [29, 82]

$a_1$: 0.0442

$a_2$: 0.3942

$a_3$: 0.5147

$b$: 37.45

Fig. 3. Scatter plot of quiz correct rate and exam score.

### D. Descriptive Statistics of Questionnaires

To understand learners' subjective perceptions of QuizGPT, this study designed two questionnaires focusing on their experiences with the "QuizGPT Platform" and the "ChatGPT Interface," respectively. Each questionnaire was divided into five aspects: accuracy, knowledge learning, attitudinal and emotional response, method and data collection, and technological acceptance. The participants were categorized into high-usage and low-usage groups based on their usage of Quiz and ChatGPT for analysis, with results presented in TABLE I.

Firstly, the high-usage group of Quiz was assigned lower scores for the accuracy of Quiz, which means that the correctness of GAI-based quizzes needs enhancement. Future work should include the development of verification algorithms or the use of other LLMs for quiz generation. Secondly, the high-usage Quiz group scored higher in learning attitude, emotion, and data collection. This suggests that the GAI-based quiz practice method effectively boosts learners' emotional engagement, increases learning interest, and reduces learning stress. Thirdly, the high-usage group of ChatGPT scored higher in all five aspects, indicating that learners find the interactive Q&A interface of ChatGPT beneficial in enhancing learning interest, easing learning stress, and making learning more enjoyable. Learners also strongly agreed with statements like "The ChatGPT assisted interface encourages me to adopt multiple learning approaches" and "I believe this 'learning through chatting' concept will be a primary tool in my future learning."

TABLE I. QUESTIONNAIRES ANALYSIS RESULT

| Category / Orientation | QuizGPT Platform | | ChatGPT Interface | |
|---|---|---|---|---|
| | Mean | | Mean | |
| | Low-Usage Group, n=20 | High-Usage Group, n=20 | Low-Usage Group, n=20 | High-Usage Group, n=20 |
| A. Accuracy | 3.90 | 3.67 | 3.66 | 4.04 |
| B. Knowledge learning | 3.95 | 3.99 | 3.71 | 4.05 |
| C. Attitudinal and emotional response | 3.68 | 3.84 | 3.49 | 3.81 |
| D. Method and data collection | 3.80 | 4.00 | 3.77 | 4.04 |
| E. Technological acceptance | 4.09 | 4.10 | 3.86 | 4.19 |

## VI. CONCLUSION

We designed an interactive interface combining Quiz with ChatGPT and employed GAI-based quizzes as a supplementary tool for learning Python. Usage data and student performance indicate that even GAI-based quizzes can generate positive educational outcomes. This study also found that students benefit from using the ChatGPT to learn Python programming. Machine learning models confirm that quiz activity parameters, including students' Quiz and ChatGPT usage, could predict students' test performance with a 3.58% Root Mean Squared Error percentage (%RMSE). Questionnaire analysis revealed that GAI-based quizzes can enhance learners' interest and reduce learning stress. However, some quizzes contained errors, leading students with high Quiz usage to rate the accuracy aspect lower scores. This phenomenon highlights the accuracy of QuizGPT should be improve in the future. In the future, error-free quiz questions could be generated by optimized LLM models, to help students learn Python programming more effectively.

## REFERENCES


[1] M. Abdullah, A. Madain, and Y. Jararweh, "ChatGPT: Fundamentals, applications and social impacts," in *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, IEEE, 2022, pp. 1–8.

[2] T. Wu *et al.*, "A brief overview of ChatGPT: The history, status quo and potential future development," *IEEECAA J. Autom. Sin.*, vol. 10, no. 5, pp. 1122–1136, 2023.

[3] I. Adeshola and A. P. Adepoju, "The opportunities and challenges of ChatGPT in education," *Interact. Learn. Environ.*, pp. 1–14, 2023.

[4] M. Sharples, "Automated essay writing: An AIED opinion," *Int. J. Artif. Intell. Educ.*, vol. 32, no. 4, pp. 1119–1126, 2022.

[5] J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?," *J. Appl. Learn. Teach.*, vol. 6, no. 1, 2023.

[6] H. Lee, "The rise of ChatGPT: Exploring its potential in medical education," *Anat. Sci. Educ.*, 2023.

[7] C. K. Boscardin, B. Gin, P. B. Golde, and K. E. Hauer, "ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity," *Acad. Med.*, vol. 99, no. 1, p. 22, Jan. 2024, doi: 10.1097/ACM.0000000000005439.

[8] U. Lee *et al.*, "Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education," *Educ. Inf. Technol.*, pp. 1–33, 2023.

[9] J. A. Hajj and M. Sah, "Assessing the Impact of ChatGPT in a PHP Programming Course," in *2023 7th International Symposium on Innovative Approaches in Smart Technologies (ISAS)*, Nov. 2023, pp. 1–10. doi: 10.1109/ISAS60782.2023.10391549.