

Chatgpt Assisted Teachers in Improving Formative Assessment

Shuting Zhang

Art and Social Science, The University of Sydney, Sydney, Australia

Szha9207@uni.sydney.edu.au

Abstract. In recent years, the application of artificial intelligence technology in education has received widespread attention, but its specific utility and challenges in formative assessment remain underexplored. This research paper investigates the value of ChatGPT in aiding teachers with exam design and providing student feedback. Through comprehensive interviews, the study analyses the actual effectiveness and challenges of using ChatGPT to enhance teaching efficiency and optimize student feedback. The findings indicate that ChatGPT offers significant advantages in generating diverse assessment questions and personalized feedback. However, teachers must validate and edit the AI-generated content to ensure its accuracy and relevance. Based on these insights, the paper provides several key recommendations: validating AI-generated content, enhancing teacher training, and maintaining teachers' professional independence. These steps are essential to optimize the quality and efficiency of educational assessments. This research aims to contribute to the better integration of AI technologies in education by addressing both their potential and their limitations comprehensively.

Keywords: ChatGPT, formative assessment, educational technology, teacher training, AI application.

1. Introduction

As AI technology continues to be explored in education, ChatGPT is an emerging tool that demonstrates potential applications in test design and student feedback. Formative assessment is a form of assessment designed to continuously assess student progress and provide immediate feedback to help students better understand and grasp what is being learned. Current research has shown that formative assessment has a significant effect on improving students' academic achievement and learning motivation. The study also found that formative assessment can effectively promote students' self-regulated learning and deep learning. However, teachers are faced with the challenge of designing high-quality assessment tasks and providing timely feedback when implementing formative assessment. In this context, this paper will explore the initial application of ChatGPT in formative assessment and analyze how it can help teachers improve teaching efficiency and optimize the student feedback process.

Through interviews with teachers, this study not only assesses the practical utility of ChatGPT but also focuses on the challenges and limitations of its application in educational assessment. This paper aims to provide strategies and insights for the rational utilization of this emerging technology in the field of education and to provide references for the optimization and innovation of educational assessment in the future.

2. Literature Review

2.1. Formative Evaluation

Generally speaking, the purpose of the test can be divided into three kinds: first, to help teachers judge whether students have reached a certain level of achievement. Second, it has an impact on students' learning, so that students can keep active in the learning content. Finally, passing an exam can reflect the success of a teacher's decision-making in teaching, such as whether a new teaching method is effective [1]. Bloom et al. defined formative and summative assessments in detail in the

1970s [2]. They pointed out that summative assessment is more focused on the assessment of teaching or learning effectiveness, so the final score is very important. To ensure the effectiveness of the assessment, the assessment covers a wider range and needs to be able to achieve the purpose of testing a wide range of abilities, but it is difficult to find the mastery of students' specific learning units through summative assessment.

Different from summative assessment feedback, formative assessment is more frequent than summative assessment. Each teaching unit is tested during the course to help learners recognize their shortcomings in a certain part and help learners learn and grow better, for example, such as classroom tests, project assignments, student self-assessment, and mutual assessment. So, the coverage is much narrower and there's no way to measure the full range of student outcomes. Sadler proposed that formative evaluation would have a direct impact on students, while summative evaluation would have more impact on students' far-reaching decisions [3].

When the purpose of the examination is to help students improve and promote knowledge acquisition, formative assessment can better meet this demand [4,5]. Formative evaluation is a very important element in the educational process. It runs through the entire educational process, constantly monitors students' learning status, slows down students' forgetting of knowledge, and provides basic support for teachers to adjust teaching strategies in real time [6]. Thus, formative assessment is designed to ensure the effectiveness of teaching activities, place students at the center of the classroom, promote self-regulated learning, and make them active participants in their educational process [7].

For this purpose, formative assessments are mostly low-stakes tests in which test scores do not have serious consequences for a student's future. Its focus is not to evaluate the overall ability of students but to provide learning direction for phased learning according to students' performance, timely adjustment of learning strategies, and so on.

2.2. Feasibility of ChatGPT Assistance

In addition, an example of combining theory and practice is provided through the case of assisting teachers in designing formative assessments with ChatGPT and its application in classroom feedback. According to Lee's research, ChatGPT can help teachers provide multiple choice, True/false, and fill-in-the-blank questions, although 42 of the 90 test questions provided by ChatGPT are unsuitable because the answers are too vague. However, ChatGPT can be used in designing formative assessment problems, but its maximum effectiveness depends on a certain degree of manual intervention [8]. Editing and selecting AI-generated questions is required, a step that ensures that the assessment questions used accurately measure the student's understanding and engagement with the content.

Nevertheless, the potential advantages of using ChatGPT in design evaluation tasks cannot be ignored. It can not only generate diverse questions, and enrich the dimensions and perspectives of assessment, but also reduce the cognitive burden of teachers by automating initial question generation. This approach not only saves teachers time designing the first draft but also allows them to focus more on optimizing and adjusting teaching strategies to improve teaching effectiveness and the depth of student learning.

According to research by Latif and Zhai, the use of fine-tuned ChatGPT performs well in the field of education for automated scoring, especially in handling complex written student responses [9]. These results provide a solid foundation for using ChatGPT to help teachers give feedback to students. First, ChatGPT's core strength lies in its language generation and understanding capabilities. This ability can help teachers quickly and accurately identify students' mistakes and deficiencies, provide targeted suggestions and explanations, and thus enhance the learning experience and effect.

Secondly, the use of ChatGPT for teaching feedback can significantly improve efficiency. In traditional educational Settings, teachers may need to spend a lot of time manually evaluating and responding to assignments and tests for each student, especially in highly engaged and large student

populations. ChatGPT can automate this process, reducing teachers' workload and giving them more time to focus on curriculum design and instructional interaction.

In addition, the feedback provided by ChatGPT can be instant. Immediate feedback can quickly correct students' misconceptions and skill deficiencies and prevent wrong habits from solidifying in students. In addition, personalized feedback can be tailored to the specific performance of each student, which helps to meet different learning needs and levels, driving progress for all students. However, despite the many potential advantages of using ChatGPT in educational assessment, there are still several aspects that need attention in practical applications. For example, teachers need to monitor and edit the generated feedback to ensure its accuracy and appropriateness.

3. Research Objectives and Methods

3.1. Research Design

3.1.1 Research purpose

This study aims to explore how ChatGPT can assist teachers in test design and student feedback. Traditional studies are mostly conducted from the perspective of students, examining the pros and cons of using ChatGPT in students' learning process or exploring teachers' attitudes and opinions on using ChatGPT. Instead, this study focuses on whether ChatGPT can be a teaching tool for teachers to conduct effective formative assessments and give accurate and constructive feedback to students. This research direction can not only supplement the gaps in the existing literature but also provide new perspectives and strategies for the application of educational technology. The purpose of this study is to reveal the practical effectiveness of ChatGPT in improving teaching efficiency and optimizing student feedback.

3.1.2 Research methods

Considering the nature and purpose of the study, the qualitative research method was adopted in this study. Data was collected through semi-structured in-depth interviews. This approach allows researchers to gain insight into how teachers use ChatGPT for test design and student feedback in real-world teaching activities. The interview will be aimed at teachers who have used ChatGPT and are familiar with the use of ChatGPT, to discuss their experience, feelings, and evaluation of the specific application effect of CHATGPT. In this way, the research aims to capture the specific practices and challenges of teachers using ChatGPT for formative assessment, as well as successful experiences in this process and possible room for improvement.

3.1.3 Main problems

The outline of this study focuses on the following main issues: First, the specific application of ChatGPT in test design and its effects, including how teachers use ChatGPT to design test questions, ChatGPT's performance in improving the efficiency and quality of test question design, and teachers' evaluation of the accuracy and applicability of ChatGPT design questions. Secondly, the specific application and effects of ChatGPT in the process of giving feedback to students, including how teachers use ChatGPT to analyze students' exam answers, ChatGPT's performance in providing feedback to students, and teachers' views on the accuracy and usefulness of feedback generated by ChatGPT. Thirdly, the challenges and limitations encountered in the use of ChatGPT in teaching assessment, including the specific problems encountered by teachers in the process of using ChatGPT, teachers' views on the limitations of ChatGPT in formative assessment, and what improvements are needed to improve the application effect of ChatGPT in teaching assessment. Finally, teachers' expectations and suggestions for the application of ChatGPT in future teaching and assessment. Based on this, this study mainly addresses the following two research questions:

How do teachers feel about using ChatGPT for teaching evaluation?

By using ChatGPT to analyze students' test answers, can teachers more effectively identify students' learning needs and provide personalized feedback?

3.2. Interviewee and Process

The interviewees in this study are teachers who have used ChatGPT and are familiar with the use of CHATGPT. This group was chosen to ensure that respondents could provide detailed usage experiences and real feedback on the application. Interviews will be conducted in the form of online meetings to facilitate the recording and collation of data. The question design of the interview will focus on the following core aspects: First, the specific application and effect of ChatGPT in exam design. Second, the specific application and effect of ChatGPT on the feedback process of students. Third, the challenges and limitations encountered when using ChatGPT for teaching assessment. Fourth, teachers' expectations and suggestions for the application of ChatGPT in future teaching and assessment

Through these interviews, the aim is to comprehensively understand the practical application of ChatGPT in teaching assessment and provide references for future research and practice.

4. Interview Results and Analysis

4.1. Benefits

Through interviews, teachers gave mixed opinions about using ChatGPT to help them design exams and give feedback. First, teachers gain significant efficiency gains in designing exam structures by using ChatGPT. Some teachers mentioned that ChatGPT can help them determine the appropriate type and number of questions according to the teaching content, to better assess the students' mastery of knowledge. Secondly, through automated analysis, ChatGPT helps teachers identify what knowledge points students are deficient in so that teachers can conduct targeted re-talk and guidance. This approach not only improves teaching efficiency but also optimizes students' learning experiences. In addition, ChatGPT can quickly provide a variety of options when generating questions and feedback, helping teachers assess student learning more comprehensively. Some teachers also mentioned that the use of ChatGPT allows them to process large volumes of student assignments and test answers more quickly, leaving them with more time to focus on tutoring individual students. Nevertheless, teachers generally agree that ChatGPT has greatly reduced their workload, reduced repetitive tasks and enabled them to focus more on teaching content and student interaction.

4.2. Existing Problems

4.2.1 Creating dependencies

However, teachers are wary of using ChatGPT directly to set questions due to concerns about possible errors, both in the questions themselves and in the accuracy of the answers. Some teachers mentioned that using ChatGPT to help analyze students' answers could help them to identify what knowledge points students were deficient in to a large extent, but it could not change the nature of teachers' understanding of students. As one teacher put it, "There was very little change in the student's understanding, and it was easy to understand but using ChatGPT to help me analyze the students' answers helped me a lot to analyze the part of the knowledge that the students did not understand clearly." In general, teachers are cautious about the use of ChatGPT, believing that it helps identify students' knowledge deficiencies, but its role in changing the nature of students' understanding is limited.

4.2.2 Increasing workload

Two teachers also mentioned their concerns about the possibility of providing false or incorrect information when using ChatGPT. Although ChatGPT is helpful in several ways, its occasional inaccuracies require teachers to be vigilant when using it and not rely solely on the data it provides. This requires teachers to maintain both critical thinking and independent judgment when using AI tools. Teachers said, "If I need to check the content generated by ChatGPT while doing my teaching, it will increase my workload." If you rely too much on AI tools like this, you may not be able to

ensure that the content is correct." While ChatGPT does not replace the role of teachers, teachers generally view it as a powerful aid. It is expected that ChatGPT can further reduce the daily teaching and assessment burden of teachers and improve work efficiency in the future. At the same time, there are concerns, mainly about the problems that can result from relying on technology, such as the potential erosion of accuracy and teacher professional autonomy.

Through the above analysis, it can be seen that ChatGPT has significant potential in teaching design and student assessment and can effectively improve teaching efficiency and quality. However, limitations should be noted, such as occasional inaccuracies and effects on teacher independence. Therefore, a balance needs to be found in future research and application to ensure that the subsidiarity of technology does not affect the core values of education and the professional development of teachers.

5. Countermeasures

5.1. Verify AI-Generated Content to Ensure Accuracy

To ensure that the content generated by ChatGPT is accurate and applicable, teachers should validate AI-generated content when using ChatGPT to design exam questions and provide feedback. Teachers can check the content generated by ChatGPT together with other teachers in the teaching and research group to ensure the accuracy and applicability of the topics. In addition, in the initial stage, teachers can use the content generated by ChatGPT as auxiliary material, in conjunction with teaching exam content designed by teachers themselves, to gradually evaluate and adjust the performance of AI. Teachers can try the content generated by ChatGPT on a small scale in actual teaching and adjust and optimize it based on feedback.

5.2. Strengthen Teacher Training and Digital Literacy

To realize ChatGPT's full potential, teachers need to be trained to improve their digital literacy and understanding of AI tools. Schools can provide specialized training courses to help teachers master the use of ChatGPT and understand its limitations and potential problems [10]. Training sessions can include how to effectively use ChatGPT to generate content, how to evaluate and adjust AI-generated questions and feedback, and more. This encourages teachers to continue to learn and update their knowledge and to keep abreast of the application of new technologies and tools. Schools or educational institutions can organize regular seminars and training courses to share the latest AI technology application cases and research results to help teachers constantly upgrade their skills.

5.3. Maintain the Professional Independence of Teachers

While ChatGPT can help teachers in many ways, teachers should always maintain their professional independence and not rely solely on AI tools. Teachers should maintain critical thinking when using ChatGPT and independently judge AI-generated content to ensure its accuracy and applicability. Teachers should have the ability to check the content generated by AI and not blindly accept the information provided by AI. Teachers should flexibly adjust and optimize teaching strategies according to the actual teaching situation to ensure that the learning effect and experience of students are always in a priority position. Teachers need to flexibly use the content generated by ChatGPT on the specific situation and feedback of students and make timely adjustments to ensure the achievement of teaching goals.

Through the above countermeasures, it can better deal with the problems encountered in the instructional design and evaluation process using ChatGPT and ensure that the subsidiarity of technology does not affect the core values of education and the professional development of teachers. This can not only improve teaching efficiency but also give full play to the advantages of AI technology under the premise of ensuring the quality of teaching.

6. Conclusion

In summary, ChatGPT has significant potential in educational assessment, especially in the design and implementation of formative assessment. ChatGPT can effectively help teachers design exam questions, and manage and analyze students' answers, thereby improving teaching efficiency and quality. For example, ChatGPT can automatically generate multiple question types, help teachers quickly construct tests, and provide real-time feedback based on student responses, all of which are important tools for improving teaching interaction and student learning outcomes. However, the study also revealed some challenges, particularly in terms of the reliability and accuracy of the dependent technology. Teachers have expressed concerns about the accuracy of the content generated by ChatGPT, fearing that over-reliance may affect the quality of teaching. In addition, teachers will need to spend additional time supervising and correcting AI-generated assessment materials, which may offset some of the time saved through automation. Overall, ChatGPT and similar AI tools show great potential in education, but their maximum use will depend on the active involvement of teachers and the appropriate use of the technology.

However, due to the rapid development of technology, the version of ChatGPT used in this study may soon be replaced by newer versions, and the associated features and performance may also change significantly. Future studies could be conducted on a wider sample to obtain more representative and statistically significant data. At the same time, it is also possible to explore the application effect of ChatGPT in different subject areas, to see if there are certain areas where it is better or if there are certain challenges. In addition, the research can further explore how to effectively combine the professional judgment of teachers and the technological advantages of AI to optimize the quality and efficiency of teaching assessment.

References

- [1] Newton P E. Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 2007, 14(2): 149-170.
- [2] Bloom B S. *Handbook on formative and summative evaluation of student learning*, 1971.
- [3] Sadler D R. Formative assessment and the design of instructional systems. *Instructional Science*, 1989, 18(2): 119-144.
- [4] Black P. *Testing: friend or foe?: Theory and practice of assessment and testing*. Routledge, 2002.
- [5] Harlen W. Teachers' summative practices and assessment for learning tensions and synergies. *Curriculum Journal*, 2005, 16(2): 207-223.
- [6] Rohrer D, Pashler H. Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, 2010, 39(5): 406-412.
- [7] Clark I. Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 2012, 24(2): 205-249.
- [8] Lee D. Using ChatGPT to design assessments to encourage HE students to engage with TED talks as supplementary course readings. *A Prompt Engineering Experiment*, 2024.
- [9] Latif E, Zhai X. Fine-tuning ChatGPT for automatic scoring. *Computers and Education: Artificial Intelligence*, 2024, 6: 100210.
- [10] Dempere J, Modugu K, Hesham A, et al. The impact of ChatGPT on higher education. *Frontiers in Education*. Frontiers Media SA, 2023, 8: 1206936.