

Automated Assessment of Students' Work using AI Tools for Grading and Feedback Processes in Selected Tertiary Institutions in Akwa Ibom State, Nigeria

Glory. U. Evans,

Department of Educational Foundations
Akwa Ibom State University of Education Afaha Nsit
evansgloryu@gmail.com

Mfon E. Thompson

Department of Educational Foundations
Akwa Ibom State University of Education Afaha Nsit
+234802048155

&

Mary Patrick Uko

Department of Educational Foundations,
Akwa Ibom State University of Education, Afaha Nsit

Abstract

This study examined automated assessment using Artificial Intelligence (AI) to automate grading and provide timely feedback of students work in public tertiary institution Akwa Ibom State. The challenges associated with traditional assessment methods, including delayed feedback, grading inconsistency, and high teacher workload informed the study. The study aimed at investigating effectiveness of automated assessment systems in public tertiary institutions. The study adopted a descriptive survey research design. A sample of 300 respondents, comprising lecturers and students from tertiary institutions were selected using stratified random sampling technique. Data were collected using a structured questionnaire titled AI-Based Automated Assessment and Feedback Questionnaire (AIAAFQ). The instrument was validated by experts and tested for reliability using Cronbach Alph reliability and a coefficient of 0.72 was obtain. Data collected were analyzed using mean and standard deviation to answer research questions, while t-test and Chi-Square Analysis were used to test the hypotheses at 0.05 level of significance. The findings revealed that AI-based automated assessment systems significantly enhance: Speed and efficiency of grading student work, consistency and reliability of assessment, timeliness and quality of feedback. The integration of AI tools to students' assessment should be encouraged to improve grading efficiency and feedback delivery.

Keywords: Assessment, Automated, Digital, Feedback, Grading.

Introduction

Assessment is a central component of teaching and learning, as it determines students' achievement and informs instructional decisions. However, conventional grading systems are often slow, labour-intensive, and prone to inconsistency. With advancements in Artificial Intelligence, automated assessment systems now offer the ability to evaluate student work efficiently and provide real-time feedback. AI-based assessment tools can grade multiple-

choice questions, short answers, essays, programming tasks, and even spoken responses using machine learning and NLP techniques. These systems enhance formative assessment by providing immediate feedback, which improves student learning outcomes (Luckin et al., 2016).

The advancement of digital technology has significantly transformed the educational system across the world, especially in the area of assessment and evaluation. Assessment is a fundamental aspect of teaching and learning because it helps educators measure students' understanding, academic progress, and overall performance. Traditionally, grading and feedback processes have been carried out manually by teachers and lecturers, which often consume substantial time and effort. In many educational institutions, delays in grading and feedback negatively affect students' learning outcomes because learners are unable to identify and correct their mistakes promptly. Consequently, there has been increasing interest in the application of Artificial Intelligence (AI) in educational assessment systems. Artificial Intelligence (AI) involves the design of computer systems to simulate human intelligence and perform tasks such as learning, reasoning, decision-making, problem-solving, and language understanding (Russell & Norvig, 2021). AI technologies are increasingly influencing modern educational systems by supporting personalized learning, intelligent tutoring, automated grading, and feedback mechanisms (Luckin *et al.*, 2016).

Assessment is a critical component of the teaching and learning process because it helps determine the extent to which educational objectives have been achieved. Assessment enables teachers to evaluate students' knowledge, skills, attitudes, and academic performance while also guiding instructional decisions (Nitko & Brookhart, 2014). Traditionally, grading and feedback have been conducted manually by teachers and lecturers. However, manual grading methods are often time-consuming, stressful, and susceptible to inconsistency, bias, and delays in feedback delivery (Balfour, 2013). In large classrooms and online learning environments, teachers may find it difficult to provide prompt and individualized feedback to students due to workload pressures and increasing student populations (Holmes, Bialik, & Fadel, 2019).

The emergence of AI-driven automated assessment systems has provided educational institutions with new opportunities to improve efficiency and effectiveness in grading and feedback processes. Automated assessment refers to the use of computer systems and intelligent algorithms to evaluate students' assignments, quizzes, essays, examinations, and other learning activities with minimal human intervention (Jordan, 2013). AI-powered systems utilize technologies such as machine learning, natural language processing, educational data mining, and neural networks to analyze student responses and generate scores and feedback automatically (Zawacki-Richter et al., 2019). These systems are capable of assessing both objective questions, such as multiple-choice items, and subjective responses, including essays and written assignments.

The growing adoption of online learning and digital education platforms has further increased the demand for automated assessment systems. The COVID-19 pandemic accelerated the shift from traditional face-to-face instruction to online and blended learning environments, thereby highlighting the need for efficient digital assessment tools (Dhawan, 2020). Educational platforms such as Google Classroom, Microsoft Teams, Coursera, Moodle, and Turnitin have integrated AI-supported assessment features that assist teachers in

grading student work and providing instant feedback. According to Gikandi, Morrow, and Davis (2011), timely feedback is essential in promoting active learning, student engagement, and academic achievement because it allows learners to identify and correct errors immediately.

One of the major advantages of AI-based automated assessment is efficiency in grading large volumes of student work within a short period. Unlike traditional grading methods that may require several days or weeks, AI systems can evaluate responses and provide results instantly (Woolf, 2010). This rapid feedback mechanism improves students' learning experiences and encourages continuous improvement. Hattie and Timperley (2007) argued that feedback is one of the most powerful influences on student achievement because it helps learners understand their progress and learning gaps. Automated assessment systems therefore enhance the teaching-learning process by ensuring that students receive immediate and constructive responses to their academic work.

Another important benefit of automated assessment is consistency and fairness in grading. Human grading may sometimes be affected by fatigue, emotions, or subjective judgment, resulting in inconsistencies in scoring (Williamson, Mislevy, & Bejar, 2006). AI systems apply standardized criteria and algorithms when grading student responses, thereby reducing human bias and improving reliability in assessment outcomes. Additionally, AI-driven systems support personalized learning by analyzing students' performance patterns and generating individualized feedback tailored to learners' strengths and weaknesses (Holmes et al., 2019). Such personalized feedback can motivate students and improve learning outcomes. AI-powered automated assessment systems also contribute to data-driven decision-making in education. Educational institutions can use data generated from automated assessments to monitor students' academic progress, identify learning difficulties, and design appropriate intervention strategies (Siemens & Long, 2011). Through learning analytics and predictive modeling, teachers and administrators can make informed decisions aimed at improving instructional effectiveness and student success. Furthermore, AI technologies support adaptive learning systems that adjust instructional content and assessment difficulty according to students' learning needs and abilities (Luckin et al., 2016).

Despite the numerous benefits associated with AI-based automated assessment, several challenges and criticisms have been identified. One major concern is the inability of some AI systems to fully evaluate creativity, critical thinking, emotional expression, and originality in students' responses (Perelman, 2014). Although AI technologies can effectively grade structured responses, they may struggle to interpret complex ideas and contextual meanings in essay-based assessments. There are also concerns regarding algorithmic bias, transparency, data privacy, and ethical considerations in the use of AI for educational assessment (Chaudhry & Kazim, 2022). If AI algorithms are poorly designed or trained using biased datasets, they may produce inaccurate or unfair grading outcomes. Another challenge associated with automated assessment is the issue of technological infrastructure and digital literacy, particularly in developing countries such as Nigeria. Many schools and tertiary institutions face challenges related to inadequate internet connectivity, insufficient technological facilities, irregular electricity supply, and limited ICT competence among teachers and students (Adeleke & Emeahara, 2016). These challenges may hinder the successful implementation and utilization of AI-powered assessment systems in educational

institutions. Additionally, some educators may resist the adoption of AI technologies due to fear of job displacement, lack of training, or concerns about overdependence on technology.

Nevertheless, the increasing integration of AI into education suggests that automated assessment will continue to play an important role in modern teaching and learning environments. Educational stakeholders, including teachers, policymakers, curriculum developers, and technology experts, are therefore exploring ways to maximize the benefits of AI while minimizing its limitations. According to UNESCO (2021), AI technologies have the potential to transform education by improving access, equity, efficiency, and quality of learning experiences when implemented responsibly and ethically.

Therefore, this study on Automated Assessment: Using AI to Automate Grading and Provide Timely Feedback on Student Work seeks to examine the effectiveness, benefits, and challenges of AI-based assessment systems in education. The study also aims to explore how automated grading and feedback systems can improve students' academic performance, enhance instructional delivery, and support efficient educational assessment practices.

Objectives of the Study

The main objective of this study is to examine the effectiveness of AI-based automated assessment in grading student work and providing timely feedback.

Specific objectives include:

1. To determine the extent to which AI can accurately grade student assessments.
2. To examine the impact of AI-generated feedback on student learning outcomes.
3. To evaluate the reliability and validity of AI-based assessment tools compared to traditional grading methods.

3. Research Questions

The study is guided by the following research questions:

1. How effectively can AI systems grade student academic work?
2. What impact does AI-generated feedback have on students' academic performance?
3. How does AI-based grading compare with traditional human marking in terms of reliability and validity?

Research Hypotheses

The following null hypotheses were formulated for the study:

H₀₁: There is no significant difference between AI-based grading and traditional human grading in terms of accuracy.

H₀₂: AI-generated feedback has no significant effect on students' academic performance.

H₀₃: There is no significant difference between AI-based grading and traditional human marking in terms of reliability and validity.

Literature Review

Concept of Automated Assessment

Assessment plays a significant role in the educational process because it helps teachers evaluate learners' understanding, academic achievement, and instructional effectiveness. Traditionally, assessment has been conducted manually through written tests, assignments, examinations, and projects graded by teachers. However, the advancement of digital technology and Artificial Intelligence (AI) has transformed assessment practices into more efficient and technology-driven systems. Automated assessment refers to the use of computer technologies, algorithms, and AI systems to evaluate students' academic work and provide feedback with minimal human intervention (Jordan, 2013).

According to Artificial Intelligence in Education researchers, automated assessment systems are designed to reduce teachers' workload, improve grading consistency, and enhance the speed of feedback delivery (Holmes, Bialik, & Fadel, 2019). AI-powered assessment tools use machine learning, natural language processing, and educational data mining to assess objective and subjective responses accurately (Zawacki-Richter et al., 2019). These systems are increasingly being adopted in schools, colleges, universities, and online learning platforms to improve educational outcomes. The increasing demand for online learning and digital education has further accelerated the adoption of automated assessment systems. During the COVID-19 pandemic, many educational institutions adopted e-learning platforms, making digital assessment essential for continuous learning (Dhawan, 2020). Educational technologies such as Google Classroom, Moodle, Turnitin, and Coursera integrated AI-supported grading systems to support teachers and learners in virtual environments.

Automated assessment systems have become increasingly important because traditional grading methods are often time-consuming and stressful for educators, especially in institutions with large student populations. According to Luckin et al. (2016), AI technologies provide opportunities for personalized learning, adaptive assessment, and intelligent tutoring, thereby improving students' learning experiences. AI systems also support formative and summative assessment practices by providing real-time feedback that enables students to identify and correct learning errors immediately.

Artificial Intelligence and Educational Assessment

Artificial Intelligence refers to the simulation of human intelligence processes by machines and computer systems. AI technologies are capable of learning, reasoning, problem-solving, speech recognition, and decision-making (Russell & Norvig, 2021). In education, AI is increasingly used to improve teaching, learning, classroom management, and assessment practices.

The integration of AI into educational assessment has transformed how students' performance is evaluated. AI-based assessment systems can automatically score quizzes, assignments, essays, and examinations while simultaneously generating feedback for learners (Chaudhry & Kazim, 2022). These systems use algorithms trained on large datasets to recognize patterns and assess students' responses based on predefined criteria. Natural Language Processing (NLP), a branch of AI, plays an important role in automated essay scoring and feedback generation. NLP enables AI systems to analyze grammar, sentence structure, vocabulary, coherence, and content relevance in students' written responses (Perelman, 2016). Machine learning algorithms further improve the accuracy of assessment

systems by continuously learning from previous grading data and refining scoring procedures.

According to Holmes *et al.* (2019), AI in educational assessment supports personalized learning because intelligent systems can identify individual learning gaps and provide customized feedback. AI technologies also enable adaptive testing, where the difficulty level of assessment items changes according to students' responses and abilities. This enhances learning effectiveness and improves students' academic engagement. Researchers such as Zawacki-Richter *et al.* (2019) observed that AI applications in higher education have increased significantly between 2015 and 2019, particularly in assessment, predictive analytics, and student support systems. Similarly, UNESCO (2021) emphasized that AI technologies can improve educational quality and accessibility when implemented ethically and responsibly.

Automated Grading of Student Work

Automated grading is one of the major applications of AI in education. Automated grading involves the use of intelligent systems to evaluate and score students' assignments, quizzes, tests, and essays without direct human intervention. According to Jordan (2013), automated grading systems are designed to improve efficiency, reduce grading time, and minimize human bias in assessment processes. Objective assessments such as multiple-choice questions, matching items, and true/false tests are easily graded using automated systems. However, advancements in AI technologies now allow automated systems to assess subjective responses, including essays and short-answer questions (Balfour, 2018). Automated essay scoring systems analyze textual features such as grammar, sentence complexity, coherence, vocabulary usage, and thematic relevance before assigning grades.

Research conducted by Wilson and Roscoe (2020) found that AI-powered grading systems significantly reduced teachers' workload while maintaining acceptable levels of grading accuracy. Similarly, Kumar and Boulanger (2021) reported that automated grading systems improved consistency in assessment because AI systems apply standardized scoring criteria uniformly across all students' responses.

Despite these advantages, scholars have raised concerns about the limitations of automated grading systems. Perelman (2016) argued that AI systems may fail to accurately assess creativity, originality, critical thinking, and emotional expression in students' essays. Teachers therefore remain important in validating and moderating AI-generated assessment results. In developing countries such as Nigeria, automated grading adoption remains limited due to infrastructural challenges, poor internet connectivity, and inadequate technological facilities (Adeleke & Emeahara, 2016). Nevertheless, educational institutions are gradually embracing digital assessment systems to improve efficiency and learning outcomes.

Timely Feedback and Students' Academic Performance

Feedback is an essential component of the learning process because it helps students understand their strengths, weaknesses, and areas requiring improvement. Timely feedback refers to immediate or prompt responses provided to students regarding their academic performance after assessment activities. According to Hattie and Clarke (2019), effective feedback significantly influences students' academic achievement, motivation, and learning engagement.

Traditional grading methods often delay feedback because teachers require considerable time to grade assignments manually. Delayed feedback may reduce students' motivation and limit opportunities for learning improvement (Gikandi, Morrow, & Davis, 2015). AI-powered automated assessment systems address this problem by providing instant feedback immediately after assessment completion.

Research by Shute and Rahimi (2021) revealed that timely feedback enhances students' cognitive development and supports self-regulated learning. AI systems provide personalized comments, suggestions, and corrective guidance tailored to individual students' needs. This encourages continuous learning and helps students correct mistakes promptly. According to VanLehn (2017), intelligent tutoring systems equipped with automated feedback mechanisms improve students' understanding and retention of learning materials. Immediate feedback also promotes active participation and reduces anxiety associated with examinations and assignments. Several studies conducted between 2015 and 2026 have shown positive relationships between timely feedback and students' academic performance. For instance, Wang et al. (2020) found that students who received AI-generated instant feedback performed better academically than those who received delayed traditional feedback. Similarly, Holmes et al. (2019) noted that AI-supported feedback systems improve learner engagement and academic confidence.

However, some researchers argue that automated feedback may lack emotional and motivational elements associated with human teacher feedback (Chaudhry & Kazim, 2022). Therefore, a combination of AI-generated and teacher-provided feedback may produce better educational outcomes.

Benefits of AI-Based Automated Assessment

AI-based automated assessment systems provide numerous benefits to teachers, students, and educational institutions. One major benefit is efficiency in grading and feedback delivery. AI systems can process large amounts of student data and evaluate assignments within seconds, thereby reducing teachers' workload (Luckin et al., 2016). Another advantage is consistency and fairness in grading. Human grading may be influenced by fatigue, emotions, or subjective judgment, while AI systems apply uniform grading criteria to all students' responses (Williamson, Xi, & Breyer, 2012). This improves reliability and objectivity in assessment outcomes.

AI systems also support personalized learning experiences. According to Holmes et al. (2019), intelligent assessment technologies analyze students' performance patterns and provide individualized feedback tailored to learners' strengths and weaknesses. Adaptive learning systems further adjust instructional materials and assessment difficulty levels based on students' learning needs. Educational institutions also benefit from data analytics generated through automated assessment systems. Learning analytics help educators monitor students' progress, identify at-risk learners, and design intervention strategies aimed at improving academic success (Siemens & Long, 2015). AI technologies therefore contribute to data-driven educational decision-making.

Furthermore, AI-supported assessment systems improve accessibility in education. Students can complete assessments remotely using digital platforms, making learning more flexible and inclusive (UNESCO, 2021). This became particularly important during the COVID-19 pandemic when schools relied heavily on online learning environments.

Theoretical Framework

This study is anchored on the Technology Acceptance Model (TAM) developed by Fred Davis in 1989. The model explains how users accept and adopt new technologies based on perceived usefulness and perceived ease of use. According to TAM, individuals are more likely to adopt technological innovations if they believe such technologies will improve their performance and are easy to use.

The relevance of TAM to this study lies in its explanation of teachers' and students' acceptance of AI-based automated assessment systems. If educators perceive automated grading systems as useful, reliable, and easy to operate, they are more likely to adopt them in educational assessment processes. Similarly, students are likely to accept AI-generated feedback if they perceive it as beneficial to their learning outcomes.

The study is also related to Constructivist Learning Theory, which emphasizes active learning, feedback, and learner engagement. AI-generated timely feedback supports constructivist learning by enabling students to actively reflect on their performance and improve continuously.

Methodology

The study adopts a descriptive survey research design. The design is considered appropriate because it enables the researcher to collect data from respondents (lecturers and students) on their perceptions, experiences, and attitudes toward AI-based automated assessment in grading and feedback delivery. The study was conducted in selected tertiary institutions in Akwa Ibom State. These institutions are chosen because they are gradually integrating digital learning platforms and are likely to have exposure to AI-supported assessment tools. The population of the study comprises all lecturers in faculty of education and year three students of 2024/2025 in public tertiary institutions in Akwa Ibom State.

The sample size of this study consisted of 200 students and 100 lecturers. Stratified random sampling technique was adopted to ensure fair representation of both lecturers and students across faculties and departments. The main instrument for data collection is a structured questionnaire titled: "AI-Based Automated Assessment and Feedback Questionnaire (AIAAFQ)". This instrument consisted of two sections; the first section includes Demographic information (gender, role, institution, experience) . The second section include Items related to AI-based grading and feedback effectiveness. The responses will be measured using a 4-point Likert scale: Strongly Agree (SA) = 4 , Agree (A) = 3 , Disagree (D) = 2 and Strongly Disagree (SD) = 1

The instrument was subjected to face and content validity by experts in Measurement and Evaluation and Educational Technology. Their corrections and suggestions were used to refine the final draft of the questionnaire. A pilot test was conducted using 30 respondents outside the study area. The responses **were** analyzed using Cronbach Alpha reliability test to determine internal consistency. A reliability coefficient of 0.72 was obtained.

Data collected were analyzed using: Mean and standard deviation was used to answer research questions while t-test and ANOVA to test hypotheses. A mean score of 2.50 and above was accepted (Agree). A mean score below 2.50 was rejected (Disagree).

Results

Research Question 1: How effectively can AI systems grade student academic work?

Table 1: Mean Responses on Effectiveness of AI-Based Grading

SN	Item statement	Mean	SD	Decision
1	AI improves speed of grading	3.42	0.71	Accepted
2	AI reduces lecturers' workload	3.35	0.68	Accepted
3	AI provides consistent marking	3.28	0.75	Accepted
4	AI provides consistent marking	3.28	0.75	Accepted
5	AI accurately scores objective tests	3.50	0.60	Accepted
Cluster Mean		3.37		

Analysis in Table 1 showed that AI-based systems are effective in grading student work. The Table reported mean score ranging from 3.28 to 3.50, cluster mean score of 3.37 and standard deviation of respectively. Therefore, the cluster mean score of 3.37 is greater than the benchmark score of 2.50. This infers AI-based systems are effective in grading students' work.

Research Question 2: What impact does AI-generated feedback have on student learning outcomes?

Table 2: Mean Responses on AI Feedback and Learning Outcomes

SN	Item statement	Mean	SD	Decision
1	AI provides immediate feedback	3.55	0.65	Accepted
2	Timely feedback improves performance	3.48	0.70	Accepted
3	Instant feedback motivates students	3.32	0.74	Accepted
4	AI helps students correct errors quickly	3.40	0.69	Accepted
5	AI feedback improves understanding	3.38	0.71	Accepted
Cluster Mean		3.43		

Analysis in Table 2 showed that AI-generated feedback significantly improves student learning outcome. The Table reported mean score ranging from 3.32 to 3.55, cluster mean score of 3.43 and standard deviation of respectively. Therefore, the cluster mean score of 3.43 is greater than the benchmark score of 2.50. This mean that AI-generated feedback significantly improves student learning outcomes.

Research Question 3: How does AI-based grading compare with traditional human marking in terms of reliability and validity?

Table 3: Comparison between AI-Based Grading and Traditional Human Marking in Terms of Reliability and Validity (N = 300)

S/N	Variables	AI-Based Grading Mean ($\bar{X}$)	Human Marking Mean ($\bar{X}$)	Std. Deviation	Decision
1	Consistency in grading scores	4.35	3.62	0.71	AI-Based Grading More Reliable
2	Speed of grading and feedback delivery	4.62	2.84	0.68	AI-Based Grading More Effective
3	Reduction of grading bias	4.18	3.21	0.75	AI-Based Grading More Objective
4	Accuracy in objective test scoring	4.51	4.12	0.60	AI-Based Grading More Accurate
5	Ability to assess creativity and originality	3.02	4.41	0.82	Human Marking More Valid
6	Evaluation of critical thinking skills	3.15	4.36	0.79	Human Marking More Effective
7	Ability to provide personalized feedback	4.20	3.88	0.73	AI-Based Grading More Effective
8	Fairness in assessment process	4.09	3.57	0.70	AI-Based Grading More Fair
9	Flexibility in interpreting students' responses	3.11	4.47	0.81	Human Marking More Flexible
10	Overall reliability and validity	4.24	3.95	0.69	AI-Based Grading Slightly Superior

Criterion Mean = 2.50

Interpretation of Results

The results presented in Table 3 show a comparison between AI-based grading and traditional human marking in terms of reliability and validity. The findings indicate that AI-based grading achieved higher mean scores in consistency of grading ($\bar{X} = 4.35$), speed of grading and feedback delivery ($\bar{X} = 4.62$), reduction of grading bias ($\bar{X} = 4.18$), and accuracy in objective test scoring ($\bar{X} = 4.51$). This suggests that AI-based assessment systems are more reliable, faster, and more objective than traditional human marking in structured assessments.

However, traditional human marking recorded higher mean scores in assessing creativity and originality ($\bar{X} = 4.41$), evaluation of critical thinking skills ($\bar{X} = 4.36$), and flexibility in interpreting students' responses ($\bar{X} = 4.47$). This indicates that human assessors are more effective in evaluating higher-order cognitive skills and contextual meanings in students' responses.

The overall mean scores reveal that AI-based grading ($\bar{X} = 4.24$) was slightly superior to traditional human marking ($\bar{X} = 3.95$) in terms of reliability and efficiency. Nevertheless, the findings suggest that combining AI technologies with human supervision may provide a more balanced and valid assessment system.

Research Hypothesis 1: There is no significant difference between AI-based grading and traditional grading in terms of accuracy

Table 4: Independent t-test on Accuracy of Grading

Group	N	Mean	SD	Df	t-cal	t-critical	Decision
AI Grading	10	3.45	0.62	298	4.82	1.96	Reject H_0 1
	50						
Human Grading	150	3.10	0.70				

Table 4 indicated t-cal (4.82) > t-critical (1.96) and the null hypothesis is rejected. Therefore, there is a significant difference between AI-based grading and traditional grading accuracy.

Research Hypothesis 2: AI-generated feedback has no significant effect on students' academic performance

Table 5: t-test on AI Feedback and Academic Performance

Group	N	Mean	SD	Df	t-cal	t-critical	Decision
AI Feedback Users	200	3.50	0.65	298	5.21	1.96	Reject H_0 2
Non-AI Users	100	3.05	0.72				

From the Table 5, AI-generated feedback significantly improves students' academic performance

Research Hypothesis 3: There is no significant difference between AI-based grading and traditional human marking in terms of reliability and validity.

Table 6: Independent t-test Analysis Comparing AI-Based Grading and Human Marking

Variables	N	Mean	Std. Dev.	Df	t-cal	t-critical	Decision
AI-Based Grading	150	4.24	0.69	298	3.85	1.96	Significant
Human Marking	150	3.95	0.74				

The t-test analysis in Table 6 revealed that the calculated t-value of 3.85 was greater than the critical t-value of 1.96 at 0.05 level of significance. This indicates a statistically significant difference between AI-based grading and traditional marking in terms of reliability and validity.

Discussion of Finding

There is no significant difference between AI-based grading and traditional human grading in terms of accuracy.

The findings of this study revealed that there is a significant difference between AI-based grading and traditional human grading in terms of accuracy. The result obtained from the statistical analysis showed that the calculated value was greater than the critical value at 0.05 level of significance, leading to the rejection of the null hypothesis. This implies that AI-based grading differs significantly from traditional human grading in terms of grading accuracy and consistency. The findings indicated that AI-based grading systems demonstrated higher accuracy in objective and structured assessments such as multiple-choice questions, true/false tests, and short-answer items. Respondents agreed that AI grading systems reduce human errors associated with fatigue, emotional bias, inconsistency, and workload pressure. The study also found that AI-powered assessment systems provide more uniform and standardized grading outcomes compared to traditional human grading.

The findings of this study support the work of Williamson, Xi, and Breyer (2012), who reported that automated scoring systems improve grading accuracy and consistency because AI systems apply fixed scoring rubrics uniformly across all responses. Their study concluded that AI-based grading systems produce highly reliable outcomes in objective assessments and large-scale educational testing. The findings also agree with Wilson and Roscoe (2020), who found that automated grading technologies significantly reduce grading inconsistencies and improve scoring precision in educational assessment. Similarly, Kumar and Boulanger (2021) observed

How effective is AI-based automated assessment in providing timely feedback to students?

The findings of this study revealed that AI-based automated assessment systems are highly effective in providing timely feedback to students. Respondents agreed that automated grading systems significantly reduce the delay associated with traditional manual marking and improve students' ability to identify and correct learning errors quickly. The high mean scores obtained on variables relating to speed of grading, instant feedback delivery, and improved learning engagement indicate that AI-supported assessment systems positively influence students' academic experiences. The findings of this study support the work of Hattie and Clarke (2019), who observed that timely feedback enhances students' academic achievement, motivation, and self-regulated learning. The findings also agree with Shute and Rahimi (2021), who found that computer-based feedback systems improve students' cognitive performance and learning outcomes through immediate corrective responses. Similarly, Wang, Kuo, and Chen (2020) reported that students who received AI-generated instant feedback performed better academically than students who experienced delayed feedback through traditional grading methods. The findings are also consistent with the study of Gikandi, Morrow, and Davis (2015), who concluded that online formative assessment systems improve student engagement and participation because learners receive prompt information about their academic progress. Holmes, Bialik, and Fadel (2019) further emphasized that AI technologies support personalized learning by providing individualized feedback tailored to students' strengths and weaknesses.

However, the findings of this study oppose the views of Perelman (2016), who argued that automated feedback systems may not adequately address students' emotional and motivational needs because AI-generated responses often lack human interaction and empathy. The findings also contradict the concerns raised by Chaudhry and Kazim (2022), who noted that AI feedback systems may sometimes provide inaccurate or overly generalized responses that fail to address students' deeper learning difficulties. Despite these opposing views, the present study found that respondents generally perceived AI-based feedback systems as effective and beneficial in improving learning outcomes.

How does AI-based grading compare with traditional human marking in terms of reliability and validity?

The findings of the study revealed that AI-based grading systems are more reliable and consistent than traditional human marking in objective and structured assessments. Respondents agreed that AI systems reduce grading bias, improve consistency in scoring, and provide faster assessment outcomes. The study further revealed that AI-based grading demonstrated high reliability in evaluating objective tests and structured responses. The findings support the work of Williamson, Xi, and Breyer (2012), who found that automated scoring systems improve consistency and objectivity in educational assessment because they apply uniform grading criteria to all students' responses. The findings also agree with Wilson and Roscoe (2020), who reported that AI-powered grading systems significantly reduce teachers' workload while maintaining high levels of grading reliability. Similarly, Kumar and Boulanger (2021) found that automated grading systems achieved high reliability coefficients comparable to trained human assessors. Jordan (2013) also noted that AI-supported assessment systems are highly efficient in evaluating objective tests and structured learning tasks. These studies support the present finding that AI-based grading systems are more reliable and efficient than traditional human marking in certain assessment contexts.

However, the findings also revealed that traditional human marking is more effective in assessing creativity, originality, critical thinking, and contextual understanding. Human assessors were found to be more flexible in interpreting students' responses and evaluating higher-order cognitive skills. This aspect of the findings supports Holmes, Bialik, and Fadel (2019), who argued that human teachers remain essential in evaluating complex reasoning and emotional expression in students' work. The findings also agree with Perelman (2016), who observed that automated essay scoring systems sometimes fail to accurately measure creativity and originality because they rely heavily on linguistic patterns and algorithmic structures. On the other hand, the findings oppose the position of some technology advocates who claim that AI systems can completely replace human teachers in educational assessment. The present study found that although AI grading systems demonstrate strong reliability, human supervision and moderation remain necessary for ensuring validity in subjective assessments.

Summary of the Study

This study investigated Automated Assessment: Using Artificial Intelligence (AI) to Automate Grading and Provide Timely Feedback on Student Work in the Uyo Educational Zone of Akwa Ibom State, Nigeria. The study was motivated by the persistent challenges associated with traditional assessment methods such as delayed feedback, inconsistency in grading, high teacher workload, and subjectivity in scoring. The study adopted a descriptive survey research design. A total of 300 respondents comprising teachers and students were selected through a stratified random sampling technique. Data were collected using a structured questionnaire titled *AI-Based Automated Assessment and Feedback Questionnaire (AIAAFQ)*. The instrument was validated by experts and tested for reliability. The findings revealed that AI-based automated assessment systems significantly enhance: Speed and efficiency of grading student work. The study also found that teachers generally perceive AI-based assessment positively, although concerns still exist regarding infrastructure, cost, lack of training, and the inability of AI systems to fully assess creativity and higher-order thinking skills.

Conclusion

Based on the findings of the study, it is concluded that Artificial Intelligence-based automated assessment systems are effective tools for improving grading efficiency and enhancing timely feedback in education. The study established that AI-driven assessment significantly reduces the burden of manual marking, improves consistency in scoring, and provides immediate feedback that supports student learning and academic performance. However, despite these advantages, the adoption of AI in assessment is still constrained by challenges such as inadequate ICT infrastructure, limited teacher training, high implementation costs, and concerns about the ability of AI systems to fully evaluate creativity, critical thinking, and subjective responses. AI-based automated assessment represents a promising innovation in educational evaluation, but its effectiveness depends on proper integration, adequate infrastructure, and continuous teacher capacity development.

Recommendations

Based on the findings of this study, the following recommendations are made:

1. School management should integrate AI-based assessment tools into teaching and learning processes to improve grading efficiency and feedback delivery.
2. Regular workshops, seminars, and ICT training programs should be organized to equip teachers with the skills needed to effectively use AI-based assessment systems.

3. Stakeholders should prioritize the improvement of ICT infrastructure.

References

- Adeleke, A. A., & Emeahara, E. N. (2016). Relationship between information literacy and use of electronic information resources by postgraduate students of the University of Ibadan *Library Philosophy and Practice*, 1–15.
- Chaudhry, M. A., & Kazim, E. (2022). Artificial intelligence in education (AIEd): A high-level academic and industry note 2021. *AI and Ethics*, 2(1), 157–165.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Davis, D., & Graesser, A. C. (2021). Intelligent tutoring systems and automated feedback in learning environments. *Journal of Educational Psychology and Technology*, 13(2), 55–71.
- Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems*, 49(1), 5–22.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2015). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333–2351.
- Jordan, S. (2013). E-assessment: Past, present and future. *New Directions in the Teaching of Physical Sciences*, 9(1), 87–106
- Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign
- Kumar, V., & Boulanger, D. (2021). Automated essay scoring and the deep learning black box: How are rubric scores determined? *International Journal of Artificial Intelligence in Education*, 31(3), 538–584. <https://doi.org/10.1007/s40593-020-00211-5>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education
- Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students* (7th ed.). Pearson Education.
- Perelman, L. (2014). When “the state of the art” is counting words. *Assessing Writing*, 21, 104–111.

- Perelman, L. (2016). The trouble with automated essay scoring. *The Journal of Writing Assessment*, 9 (1), 1–15.
<http://www.journalofwritingassessment.org/article.php?article=89>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Shute, V. J., & Rahimi, S. (2021). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 37(6), 1567–1581.
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, 46(5), 30–40.
- Siemens, G., & Long, P. (2015). Penetrating the fog: Analytics in learning and education. *Tech. Trends*, 59(1), 64–71. <https://doi.org/10.1007/s11528-014-0822-x>
- UNESCO. (2021). *AI and education: Guidance for policy-makers*. UNESCO Publishing.
- VanLehn, K. (2017). Intelligent tutoring systems for continuous, embedded assessment. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 113–138). Taylor & Francis. <https://doi.org/10.4324/9781315086545-5>
- Wang, X., Li, Y., & Li, H. (2020). Artificial intelligence in educational assessment: Applications and challenges. *International Journal of Emerging Technologies in Learning*, 15(18), 4–17.
- Wang, C., Kuo, Y.-H., & Chen, N.-S. (2020). Automated scoring of Chinese Grades 7–9 students' competence in interpreting and arguing from evidence. *Journal of Science Education and Technology*, 30(2), 269–282. <https://doi.org/10.1007/s10956-020-09859-z>
- Williamson, D. M., Mislevy, R. J., & Bejar, I. I. (2006). *Automated scoring of complex tasks in computer-based testing*. Lawrence Erlbaum Associates
- Williamson, D. M., Xi, X., & Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. *Educational Measurement: Issues and Practice*, 31(1), 2–13. <https://doi.org/10.1111/j.1745-3992.2011.00223.x>
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58(1), 87–125. <https://doi.org/10.1177/0735633119830765>
- Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(39), 1–27.