



# Determinants of Teachers' Acceptance of an AI-Powered Essay Assessment System: A TAM Study of Esygrade in Economics Education

Ramadhan Defitri Pratama\*<sup>1</sup>, Khresna Bayu Sangka<sup>2</sup>, Cicilia Dyah Sulistyningrum Indrawati<sup>2</sup>

<sup>1</sup>Master of Economics Education, Universitas Sebelas Maret, Indonesia

<sup>2</sup>Faculty of Teacher Training and Education, Universitas Sebelas Maret, Indonesia

E-mail: Ramadp\_24@student.uns.ac.id

Article Info	Abstract
<b>Article History</b> Received: November, 2025 Revised: April, 2026 Published: June, 2026	This study examines teachers' acceptance of <i>Esygrade.com</i> , an AI powered automated essay assessment system designed to support more efficient and consistent evaluation of students' conceptual understanding in economics education. Manual essay scoring remains highly time consuming and prone to inter rater inconsistency, prompting increased interest in Automated Essay Scoring (AES) and large language model-based assessment tools. However, despite rapid technological advancements, teacher acceptance of AI assessment systems remains underexplored. Using the Technology Acceptance Model, this study analyzes how perceived ease of use, perceived usefulness, and attitude shape teachers behavioral intention to adopt <i>Esygrade.com</i> . A quantitative research design was employed with 30 senior high school economics teachers who explored the system before completing a validated questionnaire adapted from recent AI TAM studies. Data were analyzed using Partial Least Squares Structural Equation Modeling. The results show that perceived ease of use strongly influences both perceived usefulness and attitude, positioning usability as the most critical factor in teacher's evaluations of AI supported assessment. Perceived usefulness significantly affects attitude but does not directly predict behavioral intention, while attitude emerges as the strongest determinant of intention to use the system. These findings indicate that teacher's willingness to adopt AI powered assessment tools depends not only on their functional benefits but also on positive affective evaluations and user readiness. The study provides theoretical insight into AI acceptance in assessment contexts and practical guidance for developers and institutions seeking to integrate AI based scoring systems into classroom practice.
<b>Keywords:</b> Perceived Usefulness, Attitude Toward Use, Behavioral Intention	
Doi: <a href="http://dx.doi.org/10.23960/E3J/v9.i1.9-16">http://dx.doi.org/10.23960/E3J/v9.i1.9-16</a>	

## INTRODUCTION

Advances in artificial intelligence have increasingly influenced educational assessment practices, especially in evaluating essay based tasks that demand complex reasoning and deeper conceptual understanding. In daily classroom practice, essay scoring remains one of the most time consuming and cognitively demanding responsibilities for teachers. As reported by Barshay (2024), even highly efficient teachers may spend up to twenty minutes assessing a single essay, and this workload can easily multiply to several dozen hours each week. Ireland (2025) further indicates that grading related duties are among the most significant contributors to teacher fatigue, influencing overall job satisfaction and contributing to higher attrition rates. These findings underscore that the challenges of manual essay assessment extend beyond technical limitations and are closely linked to broader concerns about teacher well-being and sustainable workload.

In response to these concerns, research on Automated Essay Scoring has grown rapidly. AES seeks to improve scoring efficiency and objectivity through natural language processing and machine learning approaches. Grecea Pasaribu et al. (2024) demonstrate that manual essay scoring is not only time intensive but also susceptible to inter-rater inconsistency, whereas automated approaches such as 1D Convolutional Neural Networks can provide more consistent and scalable scoring outcomes. In parallel, advances in Large Language Models such as GPT-3.5 and GPT-4 have expanded the possibilities for analyzing textual coherence, reasoning quality, and conceptual understanding. These developments

provide the technological foundation for Esygrade.com, an AI-powered essay assessment system designed to help teachers deliver more efficient, consistent, and informative evaluations.

Despite these technical advancements, much of the existing AES literature remains focused on model accuracy, algorithmic optimization, and computational performance. Comparatively little attention has been devoted to understanding how teachers, as the primary end users, perceive and accept such technologies. Studies on AI acceptance in education reveal important nuances. Kong et al. (2024) found that teachers' willingness to adopt AI tools is strongly influenced by their perceived ease of use and overall comfort during interaction. Valle et al. (2024) reported that perceived usefulness does not always translate directly into behavioral intention, especially when teachers' attitudes toward the technology are not yet firmly established. Other studies similarly note that the relationship between perceived usefulness and behavioral intention may weaken or become statistically insignificant when psychological mediators such as trust or attitude exert a stronger influence (Koppel et al., 2024; Xie & Liao, 2024). These insights highlight that successful adoption of AI driven assessment tools depends not only on the capabilities of the system but also on user readiness, confidence, and the broader pedagogical environment.

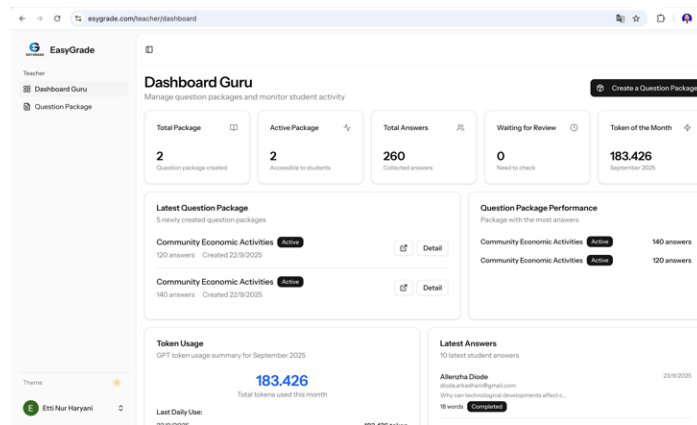
Moreover, research on AI acceptance within the specific domain of economics education remains scarce. Existing studies largely focus on general instructional technologies, tutoring systems, or chatbot-assisted learning, with limited attention to automated essay scoring using LLMs. Yet economics teachers frequently rely on essay based assessments to gauge conceptual understanding, analytical reasoning, and the ability to apply economic principles. This gap is increasingly relevant in the context of the curriculum, which emphasizes formative assessment, reflective learning, and deep conceptual engagement. While systems like Esygrade.com have the potential to support these aims, their successful integration ultimately hinges on teachers' acceptance and willingness to adopt them.

Against this backdrop, the present study examines teacher's acceptance of Esygrade.com using the Technology Acceptance Model. Specifically, the study investigates how perceived ease of use, perceived usefulness, and attitude collectively shape teachers behavioral intention to adopt an AI powered essay assessment system. The findings are expected to contribute theoretically by extending TAM to the context of LLM based AES, an area that remains underexplored. Practically, the study provides insights for developers, schools, and policymakers seeking to implement automated assessment technologies that are not only technically reliable but also aligned with teachers' expectations, needs, and readiness for innovation.

## METHODS

### System Overview

EsyGrade is an AI powered essay assessment system designed to assist teachers in evaluating students' conceptual understanding efficiently. The platform integrates Large Language Models ChatGPT to generate automated scoring and provide rubric aligned conceptual feedback. Teachers can input essay prompts and student responses through a simple interface, and the system produces both numerical scores and qualitative explanations. Figure 1 presents the EsyGrade interface used in this study.

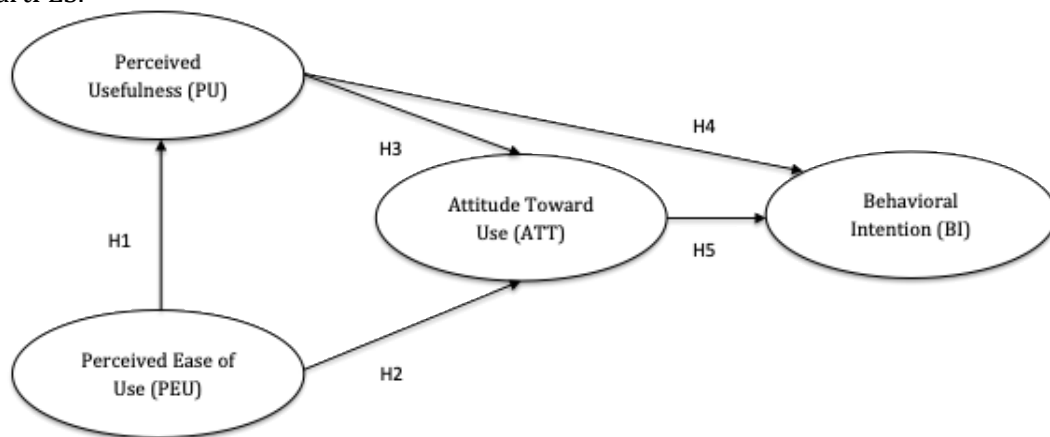


**Figure 1.** Overview of the EsyGrade platform interface used for automated essay scoring and feedback

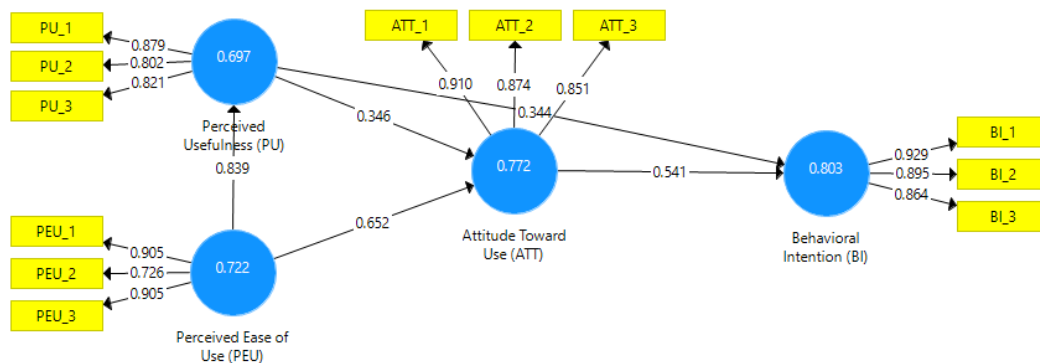
## Research Design

This study employed a quantitative research design based on the Technology Acceptance Model to examine teachers' acceptance of the EsyGrade system. A total of 30 senior high school economics teachers participated and were selected through purposive sampling. The instrument consisted of four constructs: perceived ease of use, perceived usefulness, attitude toward use, and behavioral intention. All items were adapted from Kong et al. (2024) validated TAM scales used in recent studies on AI-supported educational tools. Each item was adjusted to fit the context of automated essay assessment using EsyGrade. Responses were measured using a five-point Likert scale ranging from strongly disagree to strongly agree.

Data were collected through an online questionnaire administered after participants were introduced to EsyGrade and allowed to explore its main features. Participation was voluntary and anonymous. Data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM) with SmartPLS.



**Figure 2.** The Research Model



**Figure 3.** Algorithm Result

## RESULTS AND DISCUSSION

### A. Result

#### The outer and inner assessment model

The assessment of convergent validity showed that all indicators had loading values exceeding the recommended threshold of 0.70, indicating that they adequately represent their respective constructs. According to Hair et al., (2014), indicator loadings above 0.70 demonstrate satisfactory reliability and confirm that the indicators share sufficient variance with the underlying latent variable. The construct reliability results showed that all latent variables met the recommended reliability criteria. Cronbach's Alpha and rho\_A values exceeded the threshold of 0.70, indicating adequate internal consistency among indicators. Composite Reliability (CR) values were also above 0.70, suggesting that the constructs consistently represent their underlying dimensions. In addition, all constructs obtained Average Variance Extracted (AVE) values greater than 0.50, which confirms that more than half of the variance

in each construct is explained by its indicators. These results indicate that the constructs used in the model are valid and capable of capturing the theoretical dimensions they are intended to measure.

**Table 1.** Result Convergent Validity and Collinearity Statistics

	<b>Conbach's Alpha</b>	<b>rho_A</b>	<b>Composite Reliability</b>	<b>Average Variance Extracted (AVE)</b>
<b>PU</b>	0.782	0.792	0.873	0.697
<b>PEU</b>	0.801	0.796	0.885	0.722
<b>ATT</b>	0.852	0.855	0.910	0.772
<b>BI</b>	0.878	0.883	0.925	0.803

The collinearity assessment demonstrated that all indicators had Variance Inflation Factor (VIF) values below the upper tolerance threshold of 10, suggesting that multicollinearity is not a critical concern in the measurement model. As noted by Hair et al. (2021), VIF values below 10 are generally considered acceptable in reflective measurement models, particularly when reliability and validity criteria have been met. Therefore, the indicators were retained, and the measurement model was deemed free from problematic collinearity issues.

**Table 2.** Construct Reliability and Validity

<b>Code</b>	<b>Variable/Indicator</b>	<b>Loading</b>	<b>VIF</b>
<b>PU_1</b>	I believe that using Esygrade.com will improve my assessment performance	0.873	1.922
<b>PU_2</b>	I believe that using Esygrade.com will improve my efficiency in assessing students' essays	0.802	1.435
<b>PU_3</b>	I consider Esygrade.com to be useful for my assessment tasks	0.821	1.780
<b>PEU_1</b>	I consider learning to use Esygrade.com easy for me	0.905	3.852
<b>PEU_2</b>	I find it easy to get Esygrade.com to perform the tasks I want	0.726	1.191
<b>PEU_3</b>	I find E Esygrade.com easy to use	0.905	4.009
<b>ATT_1</b>	I believe that using Esygrade.com can enhance the quality of assessment	0.910	2.572
<b>ATT_2</b>	I like using Esygrade.com for assessment	0.874	2.051
<b>ATT_3</b>	I consider Esygrade.com to be an important tool for assessment	0.851	1.967
<b>BI_1</b>	I intend to use Esygrade.com in my assessment activities	0.929	4.120
<b>BI_2</b>	I am willing to use Esygrade.com in the future	0.895	3.429
<b>BI_3</b>	I plan to use Esygrade.com in my assessment practice	0.864	1.868

The discriminant validity assessment using the Fornell Larcker criterion showed that the square root of the AVE for each construct was greater than the correlations with other constructs. This indicates that each construct is empirically distinct and captures a unique aspect of the model. According to Hair et al. (2021), discriminant validity is established when a construct shares more variance with its own indicators than with other constructs. For additional confirmation, discriminant validity may also be assessed using the HTMT ratio, with values below 0.85–0.90 indicating adequate discriminant validity. The results confirm that all constructs in the model meet the required discriminant validity criteria.

**Table 3.** Discriminant Validity

	<b>PU</b>	<b>PEU</b>	<b>ATT</b>	<b>BI</b>
<b>PU</b>	0.853	0.839	0.893	0.827
<b>PEU</b>		0.850	0.823	0.850
<b>ATT</b>			0.879	
<b>BI</b>			0.848	0.896

### R-squared test

The R-square results show that PEU explains 70.4% of the variance in PU, indicating that ease of use contributes strongly to teachers' perceptions of the usefulness of Esygrade.com . The ATT construct obtained an R-square value of 0.923, meaning that PU and PEU together explain 92.3% of the variance in teachers' attitudes toward using the system. Meanwhile, BI recorded an R-square value of 0.744, indicating that 74.4% of teachers' intention to use Esygrade.com is explained by ATT and PU. According to Hair et al., (2014), R-square values of 0.75, 0.50, and 0.25 represent substantial, moderate, and weak levels of explanatory power. Based on these criteria, the model demonstrates substantial predictive

capability for ATT and BI, while PU shows a moderate to substantial level. Overall, these results indicate that the TAM-based model provides strong explanatory power in predicting teacher's acceptance of the ESYgrade.com system.

**Table 4. R Square Test**

	<b>R Square</b>	<b>R Square Adjusted</b>
<b>PU</b>	0.704	0.693
<b>ATT</b>	0.923	0.917
<b>BI</b>	0.744	0.725

**The size effect test (f<sup>2</sup>)**

The effect size (f<sup>2</sup>) results show the relative contribution of each exogenous construct to the endogenous variables in the model. Following Hair et al., (2014), f<sup>2</sup> values of 0.02, 0.15, and 0.35 indicate small, medium, and large effects, respectively. The analysis revealed that PEU had a very large effect on PU (f<sup>2</sup> = 2.378) and ATT (f<sup>2</sup> = 1.636), indicating that ease of use plays a dominant role in shaping both the perceived usefulness of ESYgrade.com and teachers attitudes toward using it. PU also showed a large effect on ATT (f<sup>2</sup> = 0.459), suggesting that usefulness substantially enhances teachers' positive evaluations of the system. Meanwhile, the effects of PU on BI (f<sup>2</sup> = 0.094) and ATT on BI (f<sup>2</sup> = 0.232) fall within the small to medium range, indicating that both predictors contribute meaningfully to teachers intention to use ESYgrade.com, though with different magnitudes. Overall, the effect size results demonstrate that PEU is the strongest determinant within the model, particularly in shaping PU and ATT, while both ATT and PU provide meaningful contributions to BI.

**Table 5. The Size Effect Test**

	<b>PU</b>	<b>PEU</b>	<b>ATT</b>	<b>BI</b>
<b>PU</b>			0.459	0.094
<b>PEU</b>	2.378		1.636	
<b>ATT</b>				0.232
<b>BI</b>				

**Predictive relevant test (Q<sup>2</sup>)**

The predictive relevance of the model was assessed using the Stone Geisser Q<sup>2</sup> statistic obtained through the blindfolding procedure. According to Hair et al., (2014), Q<sup>2</sup> values greater than zero indicate that the model has predictive relevance for the corresponding endogenous construct, with higher values reflecting stronger predictive accuracy. The results show that PU had a Q<sup>2</sup> value of 0.457, indicating moderate predictive relevance. ATT demonstrated a Q<sup>2</sup> value of 0.694, which reflects high predictive relevance, while BI also showed strong predictive accuracy with a Q<sup>2</sup> value of 0.567. As expected, Perceived Ease of Use yielded no Q<sup>2</sup> value because it serves as an exogenous construct in the model. Overall, these findings confirm that the model possesses substantial predictive relevance, particularly for ATT and BI, suggesting that the TAM based framework effectively predicts teacher's acceptance of the ESYgrade.com system.

**Table 6. Predictive Relevant Test Q<sup>2</sup>**

	<b>SSO</b>	<b>SSE</b>	<b>Q<sup>2</sup> (=1-SSE/SSO)</b>
<b>PU</b>	90.000	48.841	0.457
<b>PEU</b>	90.000	90.000	
<b>ATT</b>	90.000	27.550	0.694
<b>BI</b>	90.000	39.006	0.567

**The coefficient path analysis**

The hypothesis testing results show that PU had a significant positive effect on ATT ( $\beta = 0.346$ ,  $p = 0.003$ ), indicating that teachers who perceive ESYgrade.com as useful are more likely to develop favorable attitudes toward using the system. However, PU did not significantly influence BI ( $\beta = 0.344$ ,  $p = 0.147$ ), suggesting that usefulness alone is insufficient to directly drive teachers intention to use the tool. PEU demonstrated a strong and significant effect on PU ( $\beta = 0.839$ ,  $p < 0.001$ ) and ATT ( $\beta = 0.652$ ,  $p < 0.001$ ). These results indicate that ease of use is an essential determinant in shaping teachers

perceptions of usefulness and their affective evaluations of Esysgrade.com. Finally, ATT had a significant positive effect on BI ( $\beta = 0.541$ ,  $p = 0.024$ ), confirming that favorable attitudes toward Esysgrade.com meaningfully contribute to teachers willingness to adopt the system.

Following Hair et al. (2014), path coefficients are considered significant when p-values fall below the 0.05 threshold. Based on this criterion, four of the five hypothesized relationships were supported, with the exception of the path from PU to BI.

**Table 7.** Coefficient Path Analysis

	Original Sample (O)	Sample Mean (M)	Standar Deviation (STDEV)	T Statistics	P Values
PU -> ATT	0.346	0.346	0.116	2.988	0.003
PU -> BI	0.344	0.351	0.238	1.450	0.147
PEU -> PU	0.839	0.842	0.038	21.862	0.000
PEU -> ATT	0.652	0.650	0.111	5.852	0.000
ATT -> BI	0.541	0.535	0.240	2.253	0.024

## B. Discussion

The results of this study confirm that teachers' acceptance of Esysgrade.com is primarily shaped by usability and attitudinal factors rather than by perceived usefulness alone. This finding aligns with recent developments in automated essay scoring and AI based assessment, where technical performance is no longer the sole determinant of adoption. The strong effect of Perceived Ease of Use on Perceived Usefulness and Attitude indicates that teachers evaluate AI-powered assessment tools through a practical lens. This result is consistent with Kong et al. (2024), who found that teachers' acceptance of generative AI tools is driven largely by how easily the system can be integrated into existing instructional and assessment workflows. In the context of Esysgrade.com, the ability to upload essays, obtain scores, and receive conceptual feedback without complex configuration appears to reduce cognitive and operational barriers, thereby enhancing both perceived usefulness and positive attitudes. From an AES perspective, this finding reflects a broader shift in assessment technology. Recent surveys on automated essay scoring emphasize that advances in deep learning and large language models have significantly improved scoring accuracy, but adoption depends on whether such systems are perceived as manageable and trustworthy by educators (Li & Ng, 2024). Similarly, Mizumoto & Eguchi (2023) and Smerdon (2024) demonstrates that while AI language models can generate reliable essay scores, teacher acceptance remains contingent on ease of interaction and clarity of system output. The strong PEU effects observed in this study therefore suggest that Esysgrade.com design responds effectively to these practical adoption concerns.

Perceived Usefulness was found to significantly influence Attitude, indicating that teachers who recognize Esysgrade.com benefits in terms of efficiency and consistency tend to evaluate the system more positively. This result resonates with AES studies showing that automated scoring systems can reduce grading time and mitigate inter rater inconsistency, particularly for essay based assessments (Grecea Pasaribu et al., 2024; Jauhiainen et al., 2024). In economics education, where essays are frequently used to assess conceptual understanding and analytical reasoning, such advantages are pedagogically meaningful and contribute to favorable affective responses toward AI-assisted assessment. However, the nonsignificant direct relationship between Perceived Usefulness and Behavioral Intention highlights an important nuance. This pattern suggests that recognizing functional benefits does not automatically translate into adoption intention. Similar findings have been reported in TAM-based studies on educational technology, where attitude acts as a key mediating variable between usefulness and intention (Kong et al., 2024; Tang et al., 2022). In the context of AI-driven assessment, this may reflect teachers' cautious stance toward delegating evaluative authority to automated systems, even when these systems are perceived as useful.

This interpretation is further supported by research on transparency and explainability in AES. Conijn et al. (2023) show that users' trust and acceptance of automated scoring systems are strongly influenced by the presence of explanatory feedback rather than by scoring accuracy alone. Although Esysgrade.com demonstrates functional utility, teachers' intention to adopt the system appears to depend more heavily on their overall attitude, which encompasses trust, comfort, and perceived pedagogical alignment.

The significant effect of Attitude on Behavioral Intention confirms its central role in teachers' acceptance of ESYgrade.com. This finding aligns with recent TAM studies on AI adoption in education, which consistently report that positive attitudes toward AI tools are decisive predictors of usage intention (Khong et al., 2023; Kong et al., 2024). In this study, teachers who developed favorable attitudes toward ESYGrade were more willing to integrate it into their assessment practices, reinforcing the view that emotional and evaluative responses are critical in the adoption of emerging assessment technologies. Overall, the results suggest that successful implementation of AI-powered essay assessment systems requires a balance between technical capability and user-centered design. While advances in AES and large language models provide a strong technological foundation (Grecea Pasaribu et al., 2024; Li & Ng, 2024), teachers' acceptance is ultimately shaped by ease of use and positive attitudes toward the system. These findings underscore the importance of designing assessment technologies that are not only accurate and efficient but also transparent, intuitive, and aligned with teachers' pedagogical values.

## **CONCLUSIONS AND SUGGESTIONS**

### **A. Conclusion**

This study examined teachers' acceptance of ESYgrade.com, an AI-powered automated essay assessment system, using the technology acceptance model. The results demonstrate that Perceived Ease of Use is the strongest determinant influencing both Perceived Usefulness and Attitude, indicating that usability plays a central role in shaping teachers' evaluations of AI-based assessment tools. Perceived Usefulness significantly influenced Attitude but did not directly affect Behavioral Intention, suggesting that functional benefits alone are insufficient to drive adoption without positive affective responses. Attitude emerged as the key predictor of Behavioral Intention, reinforcing its mediating role in teachers' willingness to integrate AI supported assessment into their practice.

Overall, the findings highlight that successful adoption of AI-powered essay scoring systems depends not only on technological capability but also on user centered design that enhances ease of use, clarity, and perceived pedagogical alignment. ESYgrade.com predictive strength and high explanatory power suggest that AI based assessment tools have strong potential to support teachers in conducting efficient and consistent evaluations, particularly in essay heavy subjects such as economics education.

### **B. Suggestion**

Based on the findings, several practical suggestions can be offered to support the effective adoption of ESYgrade.com in educational settings. Teachers would benefit from targeted training that not only introduces the technical features of AI based assessment but also guides them in integrating such tools into their instructional and evaluative practices. Schools and institutions should provide accessible technical support and clear guidelines to help reduce uncertainty and foster user confidence, particularly given the central role of attitude in shaping intention to use. For system developers, enhancing the transparency and clarity of feedback generated by AI may strengthen trust and address the gap between perceived usefulness and behavioral intention. Future research is encouraged to involve more diverse and larger samples, incorporate additional psychological constructs such as trust and AI literacy, and explore the adoption of AI assessment tools across various subjects and educational levels to build a more comprehensive understanding of teachers' acceptance.

## **REFERENCES**

- Barshay, J. (2024). *AI Essay Grading Could Help Overburdened Teachers, But Researchers Say It Needs More Work*. The Hechinger Report. <https://www.kqed.org/mindshift/63809/ai-essay-grading-could-help-overburdened-teachers-but-researchers-say-it-needs-more-work?>
- Conijn, R., Kahr, P., & Snijders, C. (2023). The Effects of Explanations in Automated Essay Scoring Systems on Student Trust and Motivation. *Journal of Learning Analytics*, 10(1), 37–53. <https://doi.org/10.18608/jla.2023.7801>
- F. Hair Jr, J., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review*, 26(2), 106–121. <https://doi.org/10.1108/EBR-10-2013-0128>

- Grecea Pasaribu, N., Budiman, G., & Dyah Irawati, I. (2024). Auto Evaluation for Essay Assessment Using a 1D Convolutional Neural Network. *IEEE Access*, 12, 188217–188230. <https://doi.org/10.1109/ACCESS.2024.3515837>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Evaluation of Formative Measurement Models* (pp. 91–113). [https://doi.org/10.1007/978-3-030-80519-7\\_5](https://doi.org/10.1007/978-3-030-80519-7_5)
- Ireland, D. (2025). *A third of US teachers considered leaving education in last 12 months due to grading workload*. Learnosity. <https://learnosity.com/edtech-blog/a-third-of-us-teachers-considered-leaving-education-in-last-12-months-due-to-grading-workload/>
- Jauhiainen, J. S., & Garagorry Guerra, A. (2024). Generative AI in education: ChatGPT-4 in evaluating students' written responses. *Innovations in Education and Teaching International*, 1–18. <https://doi.org/10.1080/14703297.2024.2422337>
- Khong, H., Celik, I., Le, T. T. T., Lai, V. T. T., Nguyen, A., & Bui, H. (2023). Examining teachers' behavioural intention for online teaching after COVID-19 pandemic: A large-scale survey. *Education and Information Technologies*, 28(5), 5999–6026. <https://doi.org/10.1007/s10639-022-11417-6>
- Kong, S. C., Yang, Y., & Hou, C. (2024). Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. *Computers and Education: Artificial Intelligence*, 7, 100328. <https://doi.org/10.1016/j.caeai.2024.100328>
- Koppel, S., Logan, D. B., Zou, X., Kaviani, F., McDonald, H., Hair Jr, J. F., St. Louis, R. M., Molnar, L. J., & Charlton, J. L. (2024). Factors influencing behavioral intentions to use conditionally automated vehicles. *Journal of Safety Research*, 91, 423–430. <https://doi.org/10.1016/j.jsr.2024.10.006>
- Li, S., & Ng, V. (2024). Automated essay scoring: recent successes and future directions. *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*. <https://doi.org/10.24963/ijcai.2024/897>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Smerdon, D. (2024). AI in essay-based assessment: Student adoption, usage, and performance. *Computers and Education: Artificial Intelligence*, 7, 100288. <https://doi.org/10.1016/j.caeai.2024.100288>
- Tang, T. T., Nguyen, T. N., & Tran, H. T. T. (2022). Vietnamese Teachers' Acceptance to Use E-Assessment Tools in Teaching: An Empirical Study Using PLS-SEM. *Contemporary Educational Technology*, 14(3), ep375. <https://doi.org/10.30935/cedtech/12106>
- Valle, N. N., Kilat, R. V., Lim, J., General, E., Dela Cruz, J., Colina, S. J., Batican, I., & Valle, L. (2024). Modeling learners' behavioral intention toward using artificial intelligence in education. *Social Sciences & Humanities Open*, 10, 101167. <https://doi.org/10.1016/j.ssaho.2024.101167>
- Xie, S., & Liao, F. (2024). Incorporating personality traits for the study of user acceptance of electric micromobility-sharing services. *Transportation Research Part F: Traffic Psychology and Behaviour*, 107, 1015–1030. <https://doi.org/10.1016/j.trf.2024.10.023>