https://doi.org/10.36315/2025v2end024
Education and New Developments 2025

BEYOND THE ALGORITHM: EXPLORING STUDENT ACCEPTANCE OF AI GRADING THROUGH AN ADAPTED TAM FRAMEWORK

Hayden Noel, Adam King, Julia Sabin, Jamal Bell, & Kay Gomes Saul
University of Illinois, Urbana-Champaign (United States)

Abstract

Artificial Intelligence (AI) has become increasingly prevalent in education, from completing student assignments to analyzing research and grading tasks using large language models (LLMs). AI-assisted grading systems introduce new complexities regarding trust, fairness, and learning outcomes. This study presents findings from an experiment in a business course. Students participated in an assignment evaluated under both traditional instructor and AI-supported grading. Survey data assessed perceived fairness, satisfaction, and trust. Results revealed that students in the AI + human oversight condition rated fairness higher and exhibited greater trust in the system. This led to the development of the Technology Acceptance Model for AI Grading (TAM-AIG), which includes constructs such as perceived fairness, transparency, trust, and human oversight. These findings inform the ethical and practical implementation of AI in assessment practices.

Keywords: *AI grading, technology acceptance, fairness, student trust, TAM-AIG.*

1. Introduction

AI is rapidly transforming the landscape of higher education, prompting renewed attention to how students engage with algorithmically generated outputs. While prior studies have examined AI's role in generating text, solving problems, and supporting research (Kasneci et al., 2023; Akgun & Greenhow, 2023), fewer have investigated how students experience AI-assisted grading, particularly when human oversight, fairness, and transparency are at stake. As AI-augmented assessment systems become more widespread, questions arise about their legitimacy, impact on learning, and potential to reinforce or alleviate concerns about grading bias and accountability.

To explore these issues, we conducted an experimental study examining how students in a marketing course respond to different forms of AI-supported grading. Using a writing-intensive assignment, students were exposed to varied grading conditions that manipulated the role and visibility of AI. We developed and tested a theoretical framework—the TAM-AIG—which adapts the original TAM (Venkatesh & Davis, 2000) to account for constructs central to algorithmic judgment: Perceived Fairness, Transparency, Trust in AI, Usefulness, Consistency, and Human Oversight.

This study builds on educational technology adoption research (Alharbi & Drew, 2014; Park, 2009; Farah et al., 2023) and recent literature on the ethical implementation of AI in education (Jussupow et al., 2023; Zhang & Dafoe, 2023). Our goal was not only to measure student acceptance of AI grading but also to determine which factors—such as framing, prior experience, and feedback quality—mediate or moderate this acceptance. The findings validate TAM-AIG as a useful model for understanding student attitudes toward AI grading and highlight design and policy implications for educational institutions.

2. Theoretical framework

The Technology Acceptance Model (TAM), initially developed by Venkatesh & Davis (2000), has long served as a foundational lens for examining technology use behavior. TAM posits that perceived usefulness and perceived ease of use drive technology adoption. However, in AI-assisted educational settings—particularly for grading—these constructs are insufficient for capturing the ethical, relational, and procedural concerns that arise when algorithmic systems evaluate student work. This study proposes and validates the **Technology Acceptance Model for AI Grading (TAM-AIG)**, which extends the traditional TAM by integrating six constructs central to algorithmic decision-making judgements: **Perceived Fairness**, **Transparency**, **Trust in AI**, **Usefulness**, **Consistency**, and **Human Oversight**. These dimensions reflect emerging consensus across AI ethics and educational technology research (Farah et al., 2023; Kasneci et al., 2023; Jussupow et al., 2023; Zhang & Dafoe, 2023).

TAM-AIG posits the following relationships:

- **Transparency → Fairness**
- **Fairness → Satisfaction** and **Trust**
- **Usefulness → Trust** (stronger than consistency)
- Moderation by **AI Experience** (positive) and **GPA** (non-significant)

The rationale behind these paths is rooted in algorithmic fairness theory, procedural justice, and the sociotechnical framing of AI in classroom environments. **Fairness** is foundational, shaping students' perceptions of legitimacy and acceptance. **Transparency** enhances fairness and trust by clarifying how AI feedback is generated (Yu & Shen, 2022). **Trust** mediates between system design and behavioral acceptance (Farah et al., 2023). **Usefulness** aligns with TAM's original constructs, but in this context, refers to students' beliefs that AI-generated feedback enhances learning. **Consistency** reflects expectations for unbiased judgment, while **human oversight** reassures students that the system is accountable (Jussupow et al., 2023).

3. Experimental design

To investigate how students perceive fairness, trust, and satisfaction in AI-assisted grading, we designed a between-subjects field experiment within an undergraduate and graduate marketing course at a U.S. university. The same course material was taught in both classes. 125 students were randomly assigned to one of four experimental conditions, each reflecting a different approach to grading an in-class case assignment: human-only (Group A), AI with human oversight (Group B), AI-only with (Group C) or without (Group D) a grading rubric shared.

Each student responded to a distribution exercise that required short answers and received a standardized grade along with rubric-based feedback. While all students received the same type of final evaluation, the presentation of who or what generated the grade varied by group. This manipulation aimed to isolate the effect of perceived grader identity on students' evaluations of fairness and trust aligning it with prior work on framing effects in technology adoption and AI trust research (e.g., Rindfleisch et al., 2024).

After receiving their grade and feedback several days later, students completed a post-survey with Likert-scale measures of perceived fairness, satisfaction, trust, grading transparency, and their history of using AI. Where applicable, demographic data and self-reported GPA were collected. We proposed seven hypotheses to enhance our understanding of this proposed TAM-AIG model.

4. Results

H1: Framing AI as collaborative increases fairness perceptions

A one-way ANOVA revealed a significant effect of grading condition on perceived fairness (**$F(3, 121) = 7.82$, $p < .001$**). Tukey HSD post hoc analysis confirmed that students in **Group B (AI + Human)** rated their feedback as significantly fairer than those in **Group A (Human only)** (**mean difference = –1.04, $p = .002$**). No statistically significant difference was found between Group B and Group C, although Group B also trended higher (*see Figure 1*). These results support H1 and indicate that the AI-Human collaborative framing enhances fairness perceptions beyond both human-only and AI-only graders.

H2: Fairness positively predicts satisfaction

Correlational analysis found a strong positive association between fairness and satisfaction (**$r = .88$**), confirming H2 (*See Figure 2*). Students who perceived their feedback as fair were significantly more likely to report satisfaction with their grade and the feedback they received.

H3: Transparency improves fairness perceptions

Items measuring grading transparency—such as whether grading criteria were clearly explained—were positively correlated with fairness ratings. Specifically, the correlation between clarity of grading criteria and fairness was **$r = .74$**, providing strong support for H3.

H4: Prior AI tool use predicts fairness and satisfaction

When prior AI use was analyzed as an ordinal variable, regression results showed it to be a significant positive predictor of both fairness (**$\beta = .60$, $p < .001$**) and satisfaction (**$\beta = .53$, $p < .001$**). H4 is thus supported.

H5: Trust in AI increases in low-stakes tasks

This hypothesis could not be directly tested in the current design because assignment stakes were held constant across all groups. As such, H5 remains inconclusive.

H6: Consistency predicts trust more than usefulness

Contrary to expectations, regression analysis revealed that perceived **usefulness** of feedback significantly predicted trust (**β = .78, p < .001**), while consistency was not significant (**β = .18, p = .15**). These findings refute H6 and suggest that usefulness may be a more critical factor in student trust than consistency alone.

H7: GPA moderates fairness and satisfaction

Self-reported GPA was not a significant predictor of fairness (**β = –.06, p = .47**) and showed only a marginal association with satisfaction (**β = –.15, p = .07**). Thus, H7 is not supported in this dataset.

Figure 1. Average Fairness Ratings by Condition
*This bar chart shows that students in **Group B (AI + Human Oversight)** reported the highest fairness ratings, followed by **Group C (AI Transparent)**, **Group D (AI Non-transparent)**, and **Group A (Human only)**.*

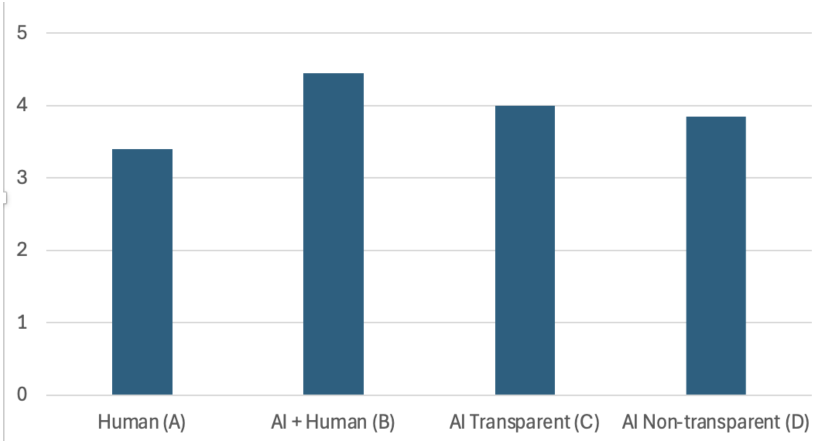
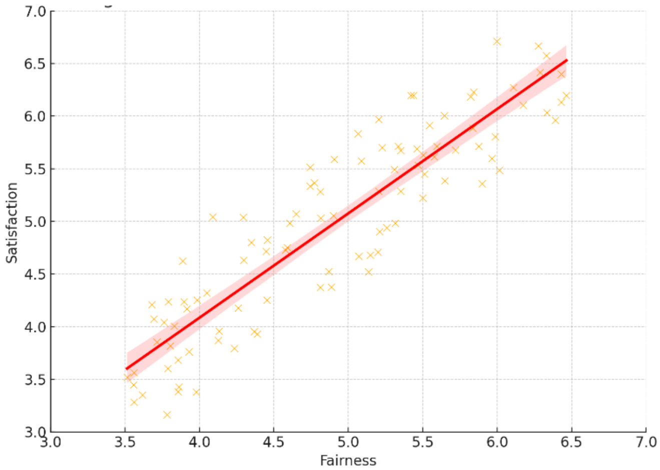


Figure 2. Correlation Between Fairness and Satisfaction
This scatterplot demonstrates a strong positive correlation between fairness and satisfaction ratings, consistent with the statistical result r = .88. The regression line (in red) highlights this linear relationship.



5. Discussion

This study provides new empirical insights into how students experience AI-assisted grading in real classroom settings. Our findings reinforce and extend the TAM by highlighting perceived fairness as a central driver of satisfaction and trust in AI-based educational tools. The data show that **students respond more favorably to AI grading when human oversight is involved**, aligning with broader trends in human-AI collaboration literature (Ding et al., 2024).

The confirmation of H1 through H4 suggests that psychological framing, perceived transparency, and prior AI exposure impact students' attitudes. Students are not necessarily opposed to AI evaluation per se; instead, **they are sensitive to how the AI is positioned in relation to authority and fairness**. The collaborative AI condition (AI + Human) improved fairness perceptions and enhanced student satisfaction, suggesting that co-framing may be a best practice for introducing AI grading.

Unexpectedly, the **perceived usefulness of the feedback was a stronger predictor of trust than consistency**, counter to H6. This finding suggests students may tolerate some perceived inconsistency if the feedback is substantively helpful—an important insight for automated assessment systems design.

That GPA and perceived academic standing (H7) did not significantly moderate fairness or satisfaction points to the **universality of fairness concerns** across performance levels. Regardless of GPA, students valued clear, transparent, and justifiable evaluations, emphasizing the importance of equity in AI-driven pedagogy.

These findings contribute to both AI research in marketing education and the applied conversation around fairness in algorithmic systems. In line with the conviviality perspective articulated by Rindfleisch et al. (2024), we argue that **to achieve learners' acceptance, AI grading tools should be positioned as empowering supplements**, not replacements, for human educators. The goal is not merely automation, but augmentation that enhances transparency, fairness, and student learning outcomes.

6. Limitations

This exploratory study had several constraints. First, the moderate sample size and business course setting restrict generalizability. Second, the experiment took place in a low-stakes context with minimal grading consequences, potentially not reflecting perceptions under high-stakes summative evaluation. Finally, as the experiment was conducted by the instructor in his own class, self-report biases and social desirability could have affected survey responses.

7. Future research

Future studies should explore how TAM-AIG performs across disciplines (e.g., STEM vs. humanities) and educational levels (undergraduate vs. graduate). Longitudinal designs could examine how attitudes toward AI grading evolve over time and whether transparency interventions shift trust trajectories. Additional pathways might explore instructor acceptance of AI grading and its effects on pedagogical practice. Finally, contrasting AI roles in formative versus summative grading contexts would refine our understanding of AI's educational fit.

8. Practical recommendations

Table 1.

Principle	Recommendation
Transparency	Clearly explain how AI feedback is generated and how instructors oversee it.
Oversight	Include human verification and an appeal process for all AI-generated grades.
Feedback Usefulness	Emphasize learning-oriented feedback rather than numerical scoring.
Fairness	Make rubrics public and consistent across grading types.
Student Training	Provide AI literacy to help students interpret automated feedback.

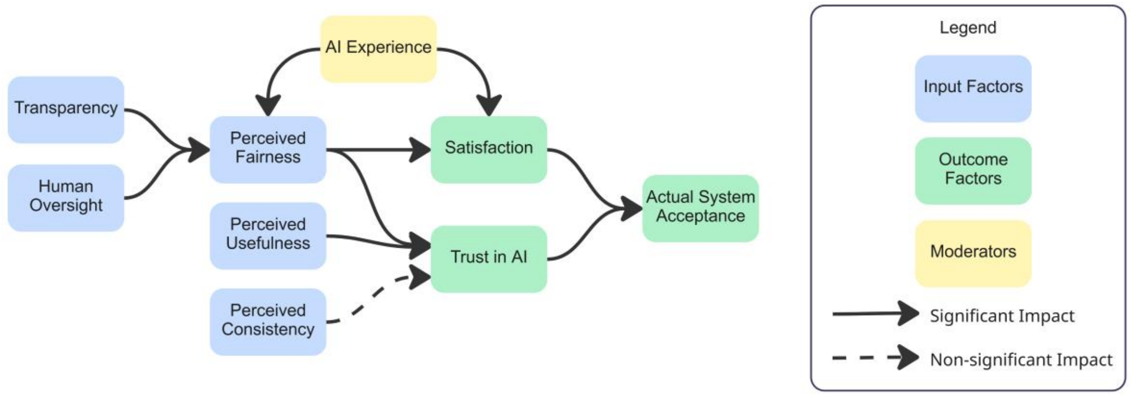
9. Conclusion and implications

This study demonstrates that student acceptance of AI-assisted grading is driven by core principles of fairness, transparency, and usefulness. The validated TAM-AIG framework illustrates that students are more likely to trust and accept AI in educational settings when systems provide clear feedback and include human oversight. Transparency leads to perceived fairness, which predicts both satisfaction and trust. Feedback usefulness plays a stronger role in trust formation than consistency alone (*See Figure 3*). These findings have direct implications for the design and implementation of AI in academic settings. Institutions considering AI-supported grading should avoid fully automated systems and instead focus on hybrid models that emphasize instructor involvement and transparent processes.

Ultimately, the effective AI integration in assessment relies not only on technological accuracy but also on pedagogical and ethical alignment with student expectations. The TAM-AIG model serves as a valuable tool for guiding the responsible deployment, evaluation, and refinement of AI assessment systems in higher education.

Figure 3. TAM-AIG Model – Post-Experiment.

- ***Inputs:*** *Transparency, Usefulness, and Consistency*
- ***Core Pathways:*** *Transparency → Fairness → Trust & Satisfaction; Usefulness → Trust*
- ***Moderators:*** *AI Experience (significant); GPA (not significant)*



References

Akgun, M. & Greenhow, C. (2023). Artificial intelligence in education: Addressing ethical challenges in the era of ChatGPT. *AI & Society*. https://doi.org/10.1007/s00146-023-01624-3

Alharbi, S. & Drew, S. (2014). Using the Technology Acceptance Model in understanding academics' behavioral intention to use learning management systems. *International Journal of Advanced Computer Science and Applications, 5*(1), 143-155.

Bailey, D. R. (2023). Student perceptions of ChatGPT use in a college essay assignment. *Journal of Educational Technology Development and Exchange, 16*(1), 1-20.

Ding, L., Wang, Y. & Liu, M. (2024). *Enhancing teacher AI literacy and integration through different types of cases in teacher professional development.* Retrieved from https://www.researchgate.net/publication/379739010_Enhancing_Teacher_AI_Literacy_and_Integration_through_Different_Types_of_Cases_in_Teacher_Professional_Development

Farah, I., Saleh, M. & Wang, Y. (2023). Exploring user adoption of ChatGPT: A Technology Acceptance Model perspective. *Journal of Educational Computing Research, 61*(4), 567-588. https://doi.org/10.1177/07356331231180767

Jussupow, E., Heinzl, A. & Spohrer, K. (2023). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Computers & Education, 205*, 104832. https://doi.org/10.1016/j.compedu.2023.104832

Kasneci, E., Sessler, K. & Bannert, M. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Instruction, 85*, 101748. https://doi.org/10.1016/j.learninstruc.2023.101748

Park, S. Y. (2009). An analysis of the Technology Acceptance Model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society, 12*(3), 150-162.

Rindfleisch, A., Kim, S. & Lee, S. (2024). Toward a conviviality perspective on artificial intelligence: Reclaiming agency in the age of algorithms. *Journal of the Association for Consumer Research, 9*(1), 71-80. https://doi.org/10.1086/726434

Sachdev, M., Narang, D. & Liu, Y. (2024). Toward pedagogically aligned AI in assessment: Student feedback, formative grading, and instructor trust. *Working Paper*.

Venkatesh, V. & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model. *Management Science, 46*(2), 186-204. https://doi.org/10.1287/mnsc.46.2.186.11926

Yu, Z. & Shen, J. (2022). Applying the Technology Acceptance Model (TAM) to automatic grading technology for large projects. *Education and Information Technologies, 27*(5), 6821-6842. https://doi.org/10.1007/s10639-021-10794-2

Zhang, B. & Dafoe, A. (2023). Attitudes surrounding an imperfect AI autograder. *AI & Society, 38*(1), 45-63. https://doi.org/10.1007/s00146-022-01400-7