Enhancing Hybrid Learning: The Role of Multimodal Adaptive Feedback in Human-AI Collaboration

AboulHassane CISSE[1,†]

[1] *Graduate School Of Informatics, Osaka Metropolitan University, Sakai, Japan*

Abstract

This study explores the transformative potential of multimodal adaptive feedback systems in hybrid learning environments, where human and artificial intelligence (AI) capabilities converge to create personalized, engaging, and effective educational experiences. By integrating real-time data from sources such as eye-tracking, emotion recognition, and physiological monitoring, these systems dynamically adapt to learners' cognitive and emotional states. The findings highlight the role of Affective Backchannels (AB) and Conversational Strategies (CS) in enhancing learner engagement, task performance, and willingness to communicate (WtC). The study demonstrates that human-AI collaboration fosters deeper learning interactions by combining human creativity, empathy, and intuition with AI's computational power and scalability. However, challenges such as scalability, diversity, and long-term applicability remain. This research underscores the need for future studies to refine these technologies, address ethical considerations, and explore broader applications in diverse educational contexts, ultimately shaping the future of inclusive and personalized learning.

Keywords

Multimodal Feedback, Hybrid Learning Environments, Human-AI Collaboration, Learner Engagement, Personalized Learning, Willingness to Communicate (WtC)

1. Introduction

The integration of Artificial Intelligence (AI) into education has transformed traditional learning environments, creating hybrid systems where human learners and intelligent systems collaborate effectively. Hybrid learning environments seek to merge human strengths such as creativity, empathy, and contextual understanding with the computational power and adaptability of AI systems. These environments represent a paradigm shift in educational practices, emphasizing personalization, adaptability, and mutual reinforcement of human-AI capabilities.

Despite significant advancements, existing research on hybrid learning has often focused on AI-driven personalization or individual components of feedback systems, without fully exploring how multimodal adaptive feedback can address cognitive and emotional gaps in learning. While prior studies have explored AI's role in education, few have systematically investigated the integration of multimodal feedback systems to bridge these gaps. This study addresses this gap by examining the role of multimodal adaptive feedback in enhancing learner engagement, task performance, and willingness to communicate (WtC).

Multimodal adaptive feedback systems, which leverage real-time data from diverse sources like eye-tracking, emotion recognition, and physiological sensors, are central to this effort. By providing dynamic, tailored responses, these systems not only improve learning outcomes but also foster deeper collaboration between humans and intelligent agents, mitigating the limitations of traditional AI approaches.

1.1 Importance of Hybrid Learning Environments

Hybrid learning environments are essential for addressing the diverse needs of learners in today's rapidly changing digital landscape. By integrating traditional teaching methods with AI-driven personalization, these environments offer inclusive and adaptive learning opportunities. Research shows that hybrid systems can bridge gaps in learner engagement and achievement, particularly in remote and technology-mediated education settings [1].

Hybrid learning environments offer **adaptability** through real-time personalized responses, **engagement** by integrating interactive elements to sustain motivation, and **collaboration** by fostering meaningful human-AI interactions that enhance mutual learning.

Joint Proceedings of LAK 2025 Workshops, co-located with the 15th International Conference on Learning Analytics and Knowledge (LAK 2025), Dublin, Ireland, March 03–07, 2025.

*Corresponding author.
† These authors contributed equally.
sn22869p@st.omu.ac.jp (C. AboulHassane) ;
0009-0009-2196-7898 (C. AboulHassane).

These systems also align with the global push toward lifelong learning, enabling individuals to acquire skills flexibly and accessibly [2]. By addressing individual differences in learning styles, preferences, and challenges, hybrid environments create an inclusive framework for modern education.

1.2 Role of Multimodal Adaptive Feedback in Human-AI Collaboration

Multimodal adaptive feedback serves as a critical tool for bridging the communication gap between humans and AI in hybrid environments. Unlike static systems, these feedback mechanisms leverage data from multiple modalities, such as eye-tracking, facial expressions, and physiological responses, to provide real-time, context-aware interactions.

Multimodal adaptive feedback systems enhance **personalization** by adjusting content in real-time based on learners' emotional states, improving focus and retention [3]. They boost **engagement** by dynamically responding to behavioral cues and stress indicators, reducing disengagement in complex tasks [4]. Additionally, they promote **collaboration** by fostering active participation, enabling hybrid systems to achieve outcomes beyond the capabilities of humans or AI alone [5].

For example, in scenarios requiring high levels of communication, such as language learning, multimodal systems can identify signs of stress or confusion through biometric data and adjust the pace or style of interaction accordingly. These systems mitigate cognitive overload and ensure that learners remain engaged, achieving outcomes beyond what either humans or AI could achieve independently.

2 Background and Related Work

2.1 Overview of Adaptive Learning

Adaptive learning systems tailor content to individual learner needs using AI and machine learning. While effective at refining learning pathways and enhancing motivation, they often fail to address emotional and cognitive challenges like managing overload. [5] highlights their potential but notes scalability and emotional adaptability as key gaps.

2.2 Multimodal Interaction Data in Education

Multimodal data such as eye-tracking, facial expressions, and physiological signals offer insights into learner behavior and engagement, enabling systems to adapt dynamically. Research like [6] demonstrates its potential in enhancing collaborative learning, yet scalability and practical integration into hybrid systems remain challenges.

2.3 Feedback Sensitivity and Personalization

Real-time, personalized feedback aligns instructional support with learners' cognitive and emotional needs. While studies such as [7] and [8] emphasize the importance of immediate, empathetic feedback, integrating these mechanisms into multimodal systems to address diverse learner needs is underexplored.

2.4 Affective Backchannels (AB) and Conversational Strategies (CS)

ABs (e.g., nods, smiles) and CSs (e.g., repetition, simplification) enhance communication and engagement in adaptive systems. [9, 11] show how combining AB and CS improves Willingness to Communicate (WtC) by reducing anxiety and building confidence. However, their integration into hybrid environments to address cognitive and emotional challenges remains limited.

3 Methodology

3.1 Experimental Design Overview

This study employed a mixed-methods research design to explore the integration of real-time gaze tracking, emotion recognition, and affective backchannels (AB) into adaptive learning environments. The objective was to evaluate the impact of these technologies on learner engagement, willingness to communicate (WtC), and overall performance. The pilot experiment was conducted in a controlled digital learning environment with approximately 9 university students participating. To ensure ethical compliance, the study followed the university's protocol guidelines for ethical approval. All participants provided signed informed consent, and the collected data were used exclusively for research purposes.

3.2 Technologies and Tools Used

Multimodal interaction data was captured using eye-tracking (Tobii Pro Nano) to assess engagement through gaze behavior, emotion recognition (OpenFace) to analyze facial expressions [10, 12], and heart rate monitoring (RookMotion) to track physiological responses like stress and relaxation [12, 13]. These tools enabled real-time biometric and emotional data analysis, supporting adaptive responses in the learning environment.

3.3 Data Collection Protocol

Participants engaged with the CeWill Conversational Agent, an AI-based system designed to integrate AB, CS and AB+CS. Data collection was structured as follows: **Pre-Test Session**: Participants completed a WtC and confidence survey to establish a baseline. **Interaction Phase**: Participants interacted with the agent, where biometric and behavioral data were collected in real-time. **Post-Test Session**: Surveys identical to the pre-test were administered, supplemented by semi-structured interviews to capture qualitative feedback on the interaction experience.

3.4 Analysis of Engagement and Emotional States

The analysis focused on understanding the relationship between multimodal data and learner outcomes: **Engagement Metrics**: Eye-tracking data such as gaze duration and fixation patterns were analyzed to measure cognitive engagement. **Emotional State Analysis**: Facial expression data were processed using Action Units (AUs) to determine predominant emotions, such as joy, frustration, or confusion, during learning activities. **Correlation Analysis**: Heart rate variability was correlated with WtC and engagement metrics to evaluate physiological responses. **Impact of Adaptive Strategies**: Comparative analysis was conducted to assess the effectiveness of AB, CS, and their combination (AB+CS) in enhancing engagement and communication.

4 Results and Findings

4.1 Enhanced Learner Engagement through Human-AI Interaction

Collaboration between humans and AI systems significantly improves learners' attention and involvement. AI systems incorporating human interaction strategies, such as affective feedback and contextual adjustments, showed a remarkable increase in engagement. **Increased Fixation Duration**: Learners demonstrated a 35% increase in time focused on key educational elements, indicating deeper immersion. **Natural Interaction**: Contextual feedback based on facial expressions and gestures created more intuitive and engaging interactions. These results affirm that embedding human-like interactions within AI systems fosters immersive and engaging educational experiences.

4.2 Improved Task Performance and Error Reduction

Human-AI collaborative systems offer superior adaptability, leading to enhanced task performance and fewer errors. Learners benefited from personalized feedback tailored to their specific mistakes. **Improved Accuracy Rates**: Task accuracy improved by 15%, supported by real-time adjustments made by the AI in response to learner actions. **Real-Time Error Correction**: Error rates decreased by 30%, attributed to precise and context-aware AI feedback. These findings illustrate how human-AI collaboration accelerates skill acquisition and ensures greater accuracy in learning tasks.

4.3 Emotional Response Management and Stress Reduction

Human-AI collaboration effectively manages learners' emotional states by integrating multimodal data such as facial expressions and physiological signals. This significantly reduced perceived stress during complex tasks. **Increased Positive Emotions**: A 20% rise in positive emotional expressions (smiles, satisfaction signals) was observed. **Reduced Stress Levels**: Heart rate variability measurements showed a notable decrease in participants' stress levels. The integration of personalized affective feedback created a safer and more motivating learning environment for learners.

4.4 Enhanced Willingness to Communicate

AI systems enriched with human strategies, such as affective backchannels (AB) and conversational strategies (CS), significantly boosted participants' WtC, especially in language learning contexts. **Increased Learner Initiatives**: The number of learner-initiated interactions increased by 25%, indicating greater confidence. **Improved Dialogue Quality**: Dynamic adjustments by conversational agents enabled smoother and more natural exchanges. These results confirm that human-AI collaboration overcomes barriers like fear of failure or lack of confidence, fostering more active communication.

5 Discussion

5.1 Implications for Hybrid Intelligence Systems

Real-time adaptive feedback systems in hybrid intelligence frameworks mark a major leap in education, dynamically adapting to learners' cognitive and emotional states using multimodal data like gaze tracking and emotion recognition. This study found a 35% engagement increase, underscoring the potential to reduce cognitive overload and enhance communication. By integrating human and AI strengths, these systems create personalized learning experiences that surpass the capabilities of traditional technologies, fostering deeper human-AI collaboration.

5.2 Scalability and Applications

Scaling these systems to diverse contexts poses challenges, including infrastructure, costs, and the need for professional training. Solutions include phased deployments, partnerships with schools and governments, and advancements in affordable tools. These approaches can expand access, making multimodal systems viable for underserved regions and promoting inclusive, lifelong learning.

5.3 Challenges and Future Research Directions

Although this study yielded promising results, it highlighted several areas requiring further exploration:

Sample Size and Diversity: Expanding studies to diverse populations ensures broader applicability.

Real-World Applications: Practical deployment requires robust infrastructure and user-friendly tools.

Longitudinal Studies: Research is needed to evaluate the long-term impact of these systems.

Ethical Considerations: Ensuring informed consent, anonymized data, and compliance with regulations like GDPR is essential.

Risk Mitigation: Prioritizing privacy, non-invasive methods, and user involvement reduces risks.

6 Conclusion

This study underscores the transformative potential of multimodal adaptive feedback systems in hybrid learning environments. By combining human strengths like creativity, empathy, and intuition with AI's computational power and adaptability, these systems create personalized, engaging, and effective educational experiences. Our findings demonstrate a 35% increase in engagement and enhanced willingness to communicate (WtC), aligning with prior studies on multimodal feedback systems [5] while extending their application to human-AI collaborative learning environments.

Despite these advancements, scalability remains a key challenge for future research and implementation. Addressing infrastructure limitations, reducing costs, and ensuring professional training are critical next steps for deploying these systems in diverse educational contexts. Phased implementations, partnerships with educational institutions, and advancements in cost-effective technologies can facilitate broader adoption.

Key Contributions:

Enhanced Learning Personalization: Adaptive systems effectively tailor content to learners' specific needs, boosting engagement and performance.

Motivation and Emotional Support: Affective responses foster a supportive, stress-reducing environment that promotes sustained learning.

Future Educational Frameworks: These systems pave the way for scalable, inclusive solutions to address global educational challenges.

Future research should validate the long-term impacts of these systems, explore their application in collaborative and interdisciplinary settings, and integrate advanced technologies to further enhance adaptability and effectiveness. Additionally, addressing ethical considerations, such as data privacy and inclusivity, will be pivotal in realizing the full potential of human-AI collaboration in education. These efforts will ensure that multimodal adaptive feedback systems become a cornerstone of future educational practices.

Acknowledgements

This work was supported by JST SPRING, Grant Number JPMJSP2138.

References

[1] Nguyen, Q., Rienties, B., & Toetenel, L. (2017). Mixing and matching learning design and learning analytics. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17), ACM, New York, NY, USA, 438–442. DOI: 10.1145/3027385.3027410

[2] Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64–71. DOI: 10.1007/s11528-014-0822-x

[3] Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. Educational Psychologist, 48(1), 25–39. DOI: 10.1080/00461520.2012.748006

[4] Çukurova, M., Luckin, R., & Kent, C. (2018). Artificial Intelligence and Multimodal Interaction in the Classroom. In Proceedings of the 2018 ACM Conference on International Computing Education Research (ICER '18), ACM, New York, NY, USA, 191–199. DOI: 10.1111/bjet.12829

[5] Gligorea, I., Cioca, M., Oancea, R., Gorski, A. -T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. Education Sciences, 13(12), 1216. https://doi.org/10.3390/educsci13121216 2.2

[6] Walkington, C., Nathan, M. J., Huang, W., & others. (2024). Multimodal analysis of interaction data from embodied education technologies. *Education Tech Research and Development, 72*, 2565–2584. https://doi.org/10.1007/s11423-023-10254-9

[7] Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

[8] Lipnevich, A. A., & Smith, J. K. (2009). Effects of differential feedback on student performance. *Journal of Educational Psychology, 101*(2), 238–255. https://doi.org/10.1037/a0014615

[9] Ayedoun, E., Hayashi, Y., & Seta, K. (2016). Web-services Based Conversational Agent to Encourage Willingness to Communicate in the EFL Context. *Journal of the Society of Instrument and Control Engineers (JSiSE), 15*(1), 15-27. https://doi.org/10.12937/ejsise.15.15

[10] T. Baltrušaitis, P. Robinson, and L. P. Morency, "OpenFace: An open source facial behavior analysis toolkit," *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, NY, USA, 2016, pp. 1–10, doi: 10.1109/WACV.2016.7477553.

[11] Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding Communicative and Affective Strategies to an Embodied Conversational Agent to Enhance Second Language Learners' Willingness to Communicate. *International Journal of Artificial Intelligence in Education, 29,* 29-57. https://doi.org/10.1007/s40593-018-0171-6

[12] AboulHassane, C., Seta, K., & Hayashi, Y. (2024). Integrative analysis of multimodal interaction data: Predicting communication dynamics and willingness to communicate (WtC) in human-agent interaction. Proceedings of LASI-Europe 2024, CEUR-WS. https://ceur-ws.org/Vol- 3738/paper4.pdf

[13] AboulHassane, C., ''Real-Time Adaptive Learning Environments Using Gaze and Emotion Recognition Engagement and Learning Outcomes'' *ICCE 2024 32nd International Conference on Computers in Education. Conference Proceedings Volume II, 726-729 (2024)*