Adaptive Feedback in Learning Environments: A Multimodal Approach to Enhancing Feedback Sensitivity and Learner Engagement

AboulHassane CISSE
sn22869p@st.omu.ac.jp
Osaka Metropolitan University
Sakai, Osaka, Japan

Abstract

This research explores the application of multimodal data, including eye-tracking, heart rate variability, and emotion recognition, to deliver real-time adaptive feedback in learning environments. Specifically, the study investigates how this feedback influences learner engagement, task performance, and persistence during English language conversations with a conversational agent in a simulated restaurant scenario. By integrating Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS), the system provides personalized feedback tailored to both emotional and cognitive states. Preliminary results from 9 participants reveal that real-time feedback effectively reduces frustration—evidenced by lower heart rates and more positive emotional expressions while significantly improving task accuracy and Willingness to Communicate (WtC). This research contributes to learning analytics and adaptive learning technologies by demonstrating how multimodal data can enhance cognitive and emotional learning outcomes. Future work will focus on expanding the dataset, refining individual physiological baselines, and exploring scalability across diverse educational settings, including more emotionally complex scenarios.

Keywords

Adaptive learning, multimodal feedback, learner engagement, emotion recognition, personalized feedback.

ACM Reference Format:
AboulHassane CISSE. 2025. Adaptive Feedback in Learning Environments: A Multimodal Approach to Enhancing Feedback Sensitivity and Learner Engagement. In *Proceedings of*
(Companion Proceedings 15th International Conference on Learning Analytics & Knowledge (LAK25)). ACM, Dublin, DUB, Ireland, 6 pages.

1 INTRODUCTION

Adaptive learning environments are designed to personalize instruction to meet individual learner needs, thereby enhancing engagement and improving outcomes. Feedback plays a pivotal role in boosting motivation and performance, particularly when delivered effectively ([Hattie and Timperley(2007)]; [Shute(2008)]). However, learners often experience cognitive overload or disengagement during complex tasks ([Sweller et al.(2011)]). Adaptive feedback, informed by both emotional and cognitive cues, has the potential to address these challenges by offering personalized, real-time support ([Azevedo and Aleven(2013)]).

While most adaptive systems emphasize cognitive performance, they frequently neglect the influence of emotional states on learning ([D'Mello and Graesser(2012)]). This study seeks to bridge that gap by leveraging multimodal data such as eye-tracking, heart rate, and emotion recognition to deliver adaptive feedback that enhances learner engagement and persistence in complex tasks. The novelty of this research lies in integrating real-time cognitive and emotional data, extending prior work on conversational dynamics with the inclusion of emotional and physiological cues ([Ayedoun et al.(2016)]; [Picard(1997)]).

In the domain of second language acquisition (SLA), Willingness to Communicate (WtC) is a critical factor in determining learners' ability to use the language effectively in real-world scenarios. Higher WtC has been linked to increased confidence, improved social interaction skills, and expanded professional opportunities in a globalized workforce ([MacIntyre et al.(1998)]). However, many learners face persistent barriers to communication due to anxiety, insufficient practice, or cultural differences.

Traditional language learning methods often fail to address these challenges, as they lack real-time, personalized feedback and opportunities for realistic dialogue. To fill this gap, advanced conversational agents have emerged as promising tools, simulating real-world interactions and providing immediate, context-sensitive feedback to help learners build confidence and communication skills.

This study explores the use of a conversational agent in a low-pressure restaurant scenario, which serves as an ideal starting point for enhancing WtC ([CISSE et al.(2024)]). Restaurant conversations are practical, familiar, and allow learners to practice structured dialogue in a controlled yet realistic setting. This environment minimizes anxiety and facilitates the development of conversational fluency. Additionally, the predictable nature of restaurant interactions makes them well-suited for personalized feedback and skill-building.

While the restaurant scenario provides an effective foundation, it has its limitations. It may be less applicable in high-stakes or emotionally charged settings, such as job interviews or public speaking events, where interactions often involve complex cultural norms

Companion Proceedings 15th International Conference on Learning Analytics & Knowledge (LAK25),
,

and emotional intelligence that are difficult to replicate computationally. Future iterations of this system could address these limitations by incorporating advanced affective computing capabilities, expanding the system's adaptability to a wider range of contexts.

2 LITERATURE REVIEW

2.1 Adaptive Learning Technologies

Adaptive learning technologies aim to personalize instruction by tailoring strategies to individual learner behavior, thereby enhancing engagement and improving outcomes. [Kulik and Fletcher(2016)] found that adaptive systems improve learning efficiency by personalizing the pace of instruction and providing targeted support. Similarly, [Desmarais and Baker(2012)] emphasized the importance of intelligent tutoring systems (ITS) in dynamically adjusting learning paths based on performance predictions. Despite their success in improving cognitive outcomes, many adaptive systems fail to leverage emotional data, underscoring the need to integrate multimodal inputs to address both cognitive and affective needs ([Azevedo and Aleven(2013)]).

2.2 Multimodal Interaction Data

The use of multimodal interaction data, such as eye-tracking, heart rate variability (HRV), and emotion recognition, provides deeper insights into learner engagement and emotional states.
[D'Mello and Graesser(2012)] demonstrated how systems like AutoTutor utilize emotional and cognitive engagement metrics to adapt feedback. [Jaques et al.(2014)] further highlighted the predictive power of HRV and emotion data in assessing engagement and task difficulty. However, few systems effectively leverage these data streams for real-time feedback adaptation, leaving a gap in the practical application of biometric inputs to personalize learning experiences.

2.3 Multimodal Learning Analytics (MMLA) and Affective-sensitive Adaptive Feedback Systems

Advancements in multimodal learning analytics (MMLA) have led to the development of systems that integrate diverse data streams to enhance learning outcomes.
For instance, [Schneider et al.(2017)] introduced the Presentation Trainer, a system that provides real-time feedback on nonverbal communication skills using multimodal data. The system's immediate and actionable feedback supports skill development during practice sessions. Building on this, [Schneider et al.(2018)] proposed the Multimodal Learning Hub (MLH), which captures and integrates customizable multimodal data configurations to support ubiquitous learning scenarios.

[Kim et al.(2018)] explored emotionally aware AI-driven smart classrooms, capable of monitoring presenters' emotional states and adjusting feedback to optimize engagement and memorability. Similarly, [Deeva et al.(2021)] reviewed automated feedback systems, highlighting the need for personalized, data-driven solutions tailored to learners' individual needs. These studies emphasize the importance of incorporating multimodal data to improve adaptive learning technologies.

Earlier systems, such as MACH (My Automated Conversation Coach) by [Hoque et al.(2013)], demonstrated the potential of leveraging multimodal data to enhance social interaction skills through real-time conversational feedback.
More recently, [Schneider et al.(2019)] extended the Presentation Trainer with an immersive Virtual Reality (VR) module, bridging the gap between practice and performance by simulating real-world scenarios.
Additionally, [Worsley(2012)] emphasized the importance of time-series analysis in MMLA, showing how temporal patterns in multimodal data can reveal learners' cognitive and emotional states, enabling more dynamic and responsive educational systems.

2.4 Feedback Sensitivity

Effective feedback plays a critical role in enhancing learning outcomes and motivation. [Shute(2008)] demonstrated that immediate feedback benefits novices, while delayed feedback fosters reflective learning. [Lipnevich and Smith(2009)] emphasized the impact of feedback tone, noting that overly critical feedback can demotivate learners. Despite these findings, many systems overlook the potential of integrating emotional and cognitive data to optimize feedback sensitivity. Incorporating biometric signals, such as HRV and emotion recognition, into feedback systems can enhance their effectiveness by making them more adaptive to individual learner needs ([D'Mello and Graesser(2012)]).

2.5 Conversational Strategies (AB, CS, and AB+CS)

Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS) have proven effective in fostering engagement and communication in intelligent tutoring systems. AB includes non-verbal cues like nodding or affirmations, which convey empathy, while CS comprises verbal prompts such as open-ended questions or clarifications ([Cassell et al.(2000)]).
[Ayedoun et al.(2016)] demonstrated that AB and CS, when combined AB+CS, significantly improve Willingness to Communicate (WtC), especially in language learning contexts. This study builds on these concepts by integrating biometric data into conversational strategies, enabling real-time, personalized feedback to enhance learner engagement and communication skills.

3 RESEARCH QUESTIONS AND HYPOTHESES

- **Research Question (RQ1):** *What are the correlations between multimodal data features (e.g., eye-tracking, heart rate, emotion) and adaptive feedback effectiveness in terms of engagement and task performance?*
  Hypothesis (H1): Multimodal data features will positively correlate with engagement and task performance, where higher emotional and cognitive cues (e.g., stable heart rate, focused eye-tracking) will indicate increased effectiveness of adaptive feedback [Shute(2008), Azevedo and Aleven(2013)].
- **Research Question (RQ2):** *How does feedback timing influence learner engagement during conversational tasks?*
  Hypothesis (H2): Real-time adaptive feedback will significantly enhance learner engagement during conversational tasks compared to delayed feedback by maintaining a steady

interaction flow and reducing frustration. [Lipnevich and Smith(2009), Sweller et al.(2011)].

- **Research Question (RQ3):** *What is the effect of feedback timing on learners' willingness to communicate (WtC) in complex scenarios?*
  Hypothesis (H3): Real-time feedback will lead to higher willingness to communicate (WtC) by improving learner confidence and persistence during complex tasks compared to delayed feedback [MacIntyre et al.(1998)].

4 METHODOLOGY

4.1 Participants

The study will involve 18–30 university students aged 18 years or older who will participate in English language conversations with a conversational agent. Participants will have basic English proficiency and primarily speak Japanese as their first language. This demographic was selected due to the well-documented challenges faced by Japanese learners in oral communication, including cultural hesitation to speak and high levels of language anxiety. These factors make this group ideal for exploring interventions aimed at improving Willingness to Communicate (WtC).

4.2 Technologies

To collect and analyze multimodal data, the experiment employed advanced technologies that provided real-time insights into participants' cognitive and emotional states. These tools ensured the adaptive feedback system was both responsive and personalized. The following technologies were utilized during the study:

- **Tobii Eye-Tracker**: This device tracked participants' gaze patterns, enabling the system to monitor attention levels and focus during conversational tasks. By analyzing fixation points and saccades, the eye-tracker identified moments of distraction or hesitation, allowing the conversational agent to provide timely corrective feedback.
- **OpenFace Software**: OpenFace, an open-source tool, was used to analyze facial expressions and detect emotional states such as frustration, confusion, or engagement. By evaluating subtle facial muscle movements, such as eyebrow raises or smiles, OpenFace captured emotional cues indicating participants' levels of comfort or difficulty during interactions. This data was essential for tailoring the agent's responses to participants' emotional needs, fostering a supportive learning environment.
- **RookMotion Device**: The RookMotion wearable measured heart rate variability (HRV), a physiological indicator of stress and cognitive load. By tracking fluctuations in HRV, the system assessed how participants responded to challenging tasks or feedback. High stress levels, indicated by reduced HRV, triggered the agent to provide simpler instructions or empathetic encouragement, ensuring participants remained engaged without feeling overwhelmed.

These technologies worked in tandem to provide a comprehensive view of participants' cognitive and emotional states during the study. The integration of eye-tracking, facial expression analysis, and HRV data ensured that feedback was context-sensitive and adaptive, enabling the conversational agent to dynamically address participants' needs. This multimodal approach allowed the system to respond effectively to real-time challenges, making the feedback more impactful and tailored to individual experiences.

4.3 Experiment Design

This experiment evaluated the effectiveness of adaptive feedback in enhancing learner engagement, task performance, and persistence. Participants engaged in simulated restaurant conversations with a conversational agent acting as a waitress, practicing conversational skills in English through tasks such as:

- Ordering food or drinks.
- Asking about menu preferences.
- Handling follow-up questions (e.g., clarifying an order).

The restaurant scenario was chosen for its practical relevance and low-stakes nature, minimizing anxiety while promoting conversational fluency. This structured context aligns with real-world scenarios and offers learners opportunities for consistent skill development, making it especially suitable for beginner and intermediate learners.

Objectives and Workflow: The study aimed to:

- Assess the impact of real-time feedback on engagement, task accuracy, and persistence.
- Explore whether the scenario promotes Willingness to Communicate (WtC).

Participants were divided into three groups based on feedback conditions. Tasks were designed to progressively evaluate participants' ability to adapt and persist through interactions.

4.4 Data Collection

4.4.1 *Quantitative Data.* Quantitative data will include:

- **Biometric Data**:
  - Gaze patterns from the Tobii eye-tracker to assess focus.
  - Emotional states from OpenFace to measure engagement or frustration.
  - HRV from the RookMotion device to track physiological stress.
- **Task Performance Metrics**: Completion rates, error rates, and conversation metrics (e.g., turns, pauses).
- **Engagement Metrics**: Time spent on tasks and frequency of feedback interactions.

To ensure accuracy, individual physiological baselines will be established for each participant before the experiment. Self-reports collected via pre- and post-surveys will triangulate biometric data and account for cultural and personal variability in emotional expression. This triangulation provides a comprehensive understanding of participants' engagement and emotional states, reducing potential misinterpretation.

4.4.2 *Qualitative Data.* Qualitative data will include:

- **Pre-Survey**:
  - Collects demographic information (e.g., English proficiency).
  - Assesses participants' confidence in using English in real-life scenarios (e.g., restaurant interactions).
- **Post-Survey**:

- Gathers feedback on participants' experiences during the conversation tasks.
- Captures perceived changes in confidence, engagement, and task difficulty.

5 DATA ANALYSIS

5.1 Quantitative Analysis

5.1.1 ***Correlation Analysis.*** Correlation analysis will be conducted to assess the relationships between biometric signals (e.g., eye-tracking, heart rate variability, emotion recognition) and task performance indicators such as completion rates and error rates. This analysis aims to determine how physiological and emotional responses influence participants' communication effectiveness.

5.1.2 ***Analysis of Variance (ANOVA).*** ANOVA will be used to compare engagement levels, task persistence, and communication performance across the three experimental conditions:

- **Real-time adaptive feedback**
- **Delayed feedback**
- **Control (non-adaptive feedback)**

This analysis will identify significant differences between conditions to evaluate the effectiveness of adaptive feedback strategies.

5.2 Qualitative Analysis

Qualitative data from pre-surveys and post-surveys will be analyzed to identify trends and changes in participants' confidence levels, engagement, and perceived usefulness of the feedback. Comparisons between pre- and post-survey responses will reveal whether the feedback influenced participants' Willingness to Communicate (WtC).

To control for multiple hypothesis testing, Bonferroni corrections will be applied to maintain robust statistical significance thresholds. This approach reduces the likelihood of false positives when examining correlations across a large number of biometric and task-related features.

5.3 Temporal Analytics

Temporal analytics will be incorporated to track how engagement metrics and emotional states evolve during each session. Inspired by [Worsley(2012)] 's work on time-series analysis in multimodal learning, the study will explore changes in:

- Heart rate variability (HRV)
- Gaze fixation patterns
- Emotional expressions

These temporal trends will provide insights into how participants adapt to feedback in real-time and how their persistence develops over successive conversational turns.

6 RESULTS

6.1 Engagement

Real-time adaptive feedback had a significant impact on learner engagement among the 9 participants. Those receiving Affective Backchannels (AB) combined with Conversational Strategies (CS) exhibited higher gaze fixation (average fixation: 0.45 for both eyes), indicating sustained attention during tasks. Additionally, participants experienced a reduction in heart rate, with the average heart rate decreasing from 81 bpm to 69 bpm during adaptive feedback sessions. This physiological change suggests that real-time feedback not only maintained engagement but also reduced stress, helping participants feel more comfortable during interactions.

Temporal analysis revealed a gradual decline in physiological stress indicators, such as heart rate, over the course of the tasks. This trend highlights increasing participant comfort with the conversational agent. Notably, Japanese learners showed significant gains in Willingness to Communicate (WtC), particularly in later stages of the interaction, demonstrating the effectiveness of adaptive feedback in reducing initial hesitation.

6.2 Task Performance

Participants who received real-time adaptive feedback demonstrated significantly higher task accuracy compared to those receiving delayed or traditional feedback. The task completion rates clearly indicate the effectiveness of integrating Affective Backchannels (AB) with Conversational Strategies (CS) to provide personalized, real-time support during learning interactions [CISSE(2024)]:

- **AB+CS group**: 92%, reflecting the benefits of immediate, adaptive feedback in maintaining focus and reducing confusion during tasks.
- **Delayed feedback group**: 74%, showing moderate improvement, but lacking the immediate corrective support needed to sustain optimal performance.
- **Control group**: 61%, emphasizing the limitations of non-adaptive feedback in supporting learners during complex tasks.

Additionally, participants in the AB+CS group exhibited the lowest error rates, further underscoring the value of real-time feedback. This reduction in errors can be attributed to the system's ability to dynamically address participants' challenges by providing context-sensitive feedback tailored to both emotional and cognitive states.

Real-time adaptive feedback enabled learners to correct mistakes more effectively and stay on track, particularly in scenarios requiring complex decision-making or multitasking. For example, during tasks involving multiple conversational turns or nuanced menu preferences, participants in the AB+CS group outperformed those in other groups by quickly adapting to suggestions provided by the conversational agent [CISSE(2024)].

This finding highlights the role of immediate feedback in reinforcing task-related behaviors, sustaining attention, and building confidence. In contrast, delayed feedback, while somewhat beneficial, failed to provide the real-time scaffolding necessary to minimize errors promptly. The control group's lower performance demonstrates the limitations of static, non-adaptive feedback in addressing real-time learning challenges.

These results reinforce the importance of integrating real-time multimodal feedback mechanisms in learning environments to optimize task performance and reduce learner frustration. Future studies could explore how this approach generalizes to more complex scenarios or higher-stakes environments [CISSE(2024)].

6.3 Emotional Responses

Emotion recognition data indicated lower frustration levels among participants during real-time adaptive feedback sessions. For example:

- **Lip corner depression (AU15_R)**: Average of 0.24 in the AB+CS group, compared to 0.36 in the control group.
- **Smile intensity (AU06_R)**: Average of 1.32 in the real-time feedback group, reflecting greater engagement and satisfaction during tasks.

While physiological data demonstrated reduced frustration, self-reports validated these findings by confirming positive emotional experiences among participants. However, discrepancies in a subset of participants emphasize the importance of considering cultural and individual variability in interpreting biometric signals.

The restaurant setting was an effective experimental context due to its low-stakes nature, which mitigated anxiety and facilitated second-language communication practice. Participants rehearsed structured dialogues in a controlled environment, providing a solid foundation for building conversational skills. Future research could expand this approach to high-stakes scenarios, such as job interviews or public speaking, to explore its applicability in emotionally charged contexts.

6.4 Learner Persistence

Learner persistence was notably higher among participants receiving real-time adaptive feedback compared to those in the control group, highlighting the impact of personalized support on sustained engagement. Persistence rates were as follows:

- **AB+CS group**: 88%, indicating that immediate and tailored feedback effectively encouraged participants to stay engaged, even during complex and demanding tasks.
- **Control group**: 65%, reflecting the challenges faced by participants without adaptive feedback in maintaining focus and perseverance during interactions.

Real-time feedback enabled participants to overcome barriers in conversational tasks, such as navigating intricate menu options or responding to unexpected follow-up questions. The adaptive nature of the feedback, which dynamically responded to participants' cognitive and emotional states, provided the necessary scaffolding to help them persist in their efforts. For example, learners in the AB+CS group reported feeling more confident and supported when faced with conversational challenges, attributing this to the agent's empathetic and contextually relevant cues.

These results suggest that adaptive feedback strategies play a pivotal role in enhancing learners' Willingness to Communicate (WtC), especially in scenarios requiring persistence and problem-solving. Participants receiving real-time feedback demonstrated greater resilience, maintaining their willingness to engage with the conversational agent despite encountering complex or unfamiliar situations. This aligns with prior findings that emphasize the role of personalized feedback in fostering persistence by reducing cognitive load and mitigating frustration [Shute(2008), Sweller et al.(2011)].

In contrast, participants in the control group often struggled to maintain engagement during difficult tasks, as static, non-adaptive feedback lacked the flexibility to address their individual needs. This resulted in higher dropout rates or incomplete conversational exchanges, further underscoring the limitations of traditional feedback mechanisms.

The findings reinforce the potential of real-time adaptive feedback to support learners in maintaining focus and motivation during extended tasks, ultimately enhancing their WtC in complex, dynamic scenarios. Future research could explore how such strategies perform in high-stakes environments, such as professional interviews or academic presentations, to assess their broader applicability.

7 DISCUSSION

This study demonstrates that real-time adaptive feedback, particularly when combining Affective Backchannels (AB) and Conversational Strategies (CS), significantly enhances engagement, task performance, and emotional responses. Tailored, immediate feedback reduced frustration and stress, as evidenced by lower heart rates and more positive emotional expressions. In contrast, delayed feedback resulted in lower engagement and task completion rates, highlighting the importance of real-time feedback for maintaining flow and confidence during communication.

By integrating emotional and physiological data (e.g., eye-tracking, heart rate, emotion recognition), this study advances learning analytics and adaptive learning technologies. The findings emphasize the need to address both cognitive and emotional dimensions in multimodal learning environments to create more personalized and effective feedback systems.

7.1 Limitations

The small sample size of 9 participants limits the generalizability of the findings. Expanding the study to include a larger, more diverse participant pool is necessary to validate the results and assess scalability. Additionally, the restaurant scenario, while effective for fostering low-stakes conversational confidence, may not reflect the complexities of high-stakes or emotionally charged environments.

7.2 Implications and Future Research

The conversational agent demonstrated efficacy in enhancing learners' Willingness to Communicate (WtC) by addressing anxiety and providing structured, personalized feedback. This approach bridges the gap between classroom instruction and real-world communication challenges, offering learners a practical, controlled setting for skill development.

However, the system's utility in high-pressure contexts, such as job interviews or public speaking, remains limited. These scenarios demand advanced conversational strategies, greater cultural sensitivity, and the ability to handle dynamic emotional responses, which the current system does not fully replicate. Future research should explore:

(1) The adaptability of conversational agents in high-stakes environments.
(2) The integration of advanced affective computing to simulate nuanced emotional and cultural interactions.
(3) The system's effectiveness with advanced learners who require more diverse and spontaneous interactions.

While the restaurant scenario proved valuable for structured practice, future iterations could extend to emotionally complex settings to better reflect the challenges learners face in professional and social interactions.

8 CONCLUSION

This study's preliminary findings demonstrate that real-time adaptive feedback significantly enhances learner engagement, task performance, emotional responses, and persistence during conversational tasks. By integrating Affective Backchannels (AB) and Conversational Strategies (CS), the system created a supportive and personalized learning environment. Participants receiving feedback based on biometric data—such as eye-tracking, heart rate, and emotion recognition—showed greater engagement and reduced stress compared to those receiving delayed or traditional feedback.

The study underscores the potential of multimodal interaction data in personalizing feedback to address both cognitive and emotional dimensions of learning. While the results are promising, the small sample size of 9 participants limits generalizability. Future research will expand the sample and explore diverse educational contexts to provide deeper insights into the long-term effects of multimodal feedback on learning outcomes.

As this research evolves, the goal is to refine adaptive feedback technologies to ensure scalability and reliability in personalized learning environments. Future efforts will focus on:

(1) Enhancing the triangulation of biometric and qualitative data for more accurate emotional state interpretation.
(2) Refining methods for establishing individual physiological baselines.
(3) Expanding self-report measures to better capture nuanced cultural and individual differences.

These advancements will contribute to the development of adaptive learning systems capable of improving educational outcomes across various fields.

Acknowledgments

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

References

[Ayedoun et al.(2016)] Emmanuel Ayedoun, Yuki Hayashi, and Kazuhisa Seta. 2016. 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* 26, 1 (2016), 1–28. https://doi.org/10.1007/s40593-018-0171-6

[Azevedo and Aleven(2013)] Roger Azevedo and Vincent Aleven (Eds.). 2013. *International Handbook of Metacognition and Learning Technologies.* Springer.

[Cassell et al.(2000)] Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth Churchill (Eds.). 2000. *Embodied Conversational Agents.* MIT Press.

[CISSE(2024)] Aboul Hassane CISSE. 2024. Real-time Adaptive Learning Environments Using Gaze and Emotion Recognition Engagement and Learning Outcomes. *32nd International Conference on Computers in Education (ICCE 2024)* (2024).

[CISSE et al.(2024)] Aboul Hassane CISSE, Kazuhisa Seta, and Yuki Hayashi. 2024. Integrative Analysis of Multimodal Interaction Data: Predicting Communication Dynamics and Willingness to Communicate (WtC) in Human-Agent Interaction. *Learning Analytics Summer Institute Europe 2024 (LASI Europe 2024 DC)* (2024). https://ceur-ws.org/Vol-3738/paper4.pdf

[Deeva et al.(2021)] Galina Deeva, Daria Bogdanova, Estefanía Serral, Monique Snoeck, and Jochen De Weerdt. 2021. A review of automated feedback systems for learners: Classification framework, challenges, and opportunities. *Computers in Education* (2021).

[Desmarais and Baker(2012)] Michel C. Desmarais and Ryan S. J. D. Baker. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22, 1 (2012), 9–38. https://doi.org/10.1007/s11257-011-9106-8

[D'Mello and Graesser(2012)] Sidney K. D'Mello and Arthur C. Graesser. 2012. AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems* 2, 4 (2012), 23. https://doi.org/10.1145/2395123.2395128

[Hattie and Timperley(2007)] John Hattie and Helen Timperley. 2007. The power of feedback. *Review of Educational Research* 77, 1 (2007), 81–112. https://doi.org/10.3102/003465430298487

[Hoque et al.(2013)] Mohammed Ehsan Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W. Picard. 2013. MACH: My automated conversation coach. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing.*

[Jaques et al.(2014)] Natasha Jaques, Cristina Conati, Jason M. Harley, and Roger Azevedo. 2014. Predicting Affect from Gaze Data during Interaction with an Intelligent Tutoring System. In *Intelligent Tutoring Systems (Lecture Notes in Computer Science, Vol. 8474).* Springer, 29–38. https://doi.org/10.1007/978-3-319-07221-0_4

[Kim et al.(2018)] Yelin Kim, Tolga Soyata, and Reza Feyzi Behnagh. 2018. Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. *IEEE Access* (2018).

[Kulik and Fletcher(2016)] James A. Kulik and John D. Fletcher. 2016. Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research* 86, 1 (2016), 42–78. https://doi.org/10.3102/0034654315581420

[Lipnevich and Smith(2009)] Anastasiya A. Lipnevich and Jeffrey K. Smith. 2009. The effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied* 15, 4 (2009), 319–333. https://doi.org/10.1037/a0017841

[MacIntyre et al.(1998)] Peter D. MacIntyre, Richard Clément, Zoltán Dörnyei, and Kimberly A. Noels. 1998. Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal* 82, 4 (1998), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

[Picard(1997)] Rosalind W. Picard. 1997. *Affective Computing.* MIT Press.

[Schneider et al.(2017)] Jan Schneider, Dirk Börner, Peter van Rosmalen, and Marcus Specht. 2017. Presentation Trainer: What experts and computers can tell about your nonverbal communication. *Journal of Computer Assisted Learning* 33, 2 (2017), 164–177. https://doi.org/10.1111/jcal.12175

[Schneider et al.(2019)] Jan Schneider, Dirk Börner, Peter van Rosmalen, and Marcus Specht. 2019. Beyond reality: Extending a presentation trainer with an immersive virtual reality module. *Sensors* (2019).

[Schneider et al.(2018)] Jan Schneider, Daniele Di Mitri, Bibeg Limbu, and Hendrik Drachsler. 2018. Multimodal Learning Hub: A tool for capturing customizable multimodal learning experiences. In *European Conference on Technology Enhanced Learning (EC-TEL).*

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

[Sweller et al.(2011)] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. *Cognitive Load Theory.* Springer.

[Worsley(2012)] Marcelo Worsley. 2012. Multimodal learning analytics: Enabling the future of learning through multimodal data analysis and interfaces. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI '12).* 353–356. https://doi.org/10.1145/2388676.2388755