MOSAIC-F: A Framework for Enhancing Students' Oral Presentation Skills through Personalized Feedback

Alvaro **Becerra**[1], Daniel **Andres**[1,†], Pablo **Villegas**[1,†], Roberto **Daza**[2] and Ruth **Cobos**[1]

[1]*GHIA Group, School of Engineering, Universidad Autónoma de Madrid, Spain*
[2]*BiDA-Lab Group, School of Engineering, Universidad Autónoma de Madrid, Spain*

Abstract

In this article, we present a novel multimodal feedback framework called MOSAIC-F, an acronym for a data-driven Framework that integrates Multimodal Learning Analytics (MMLA), Observations, Sensors, Artificial Intelligence (AI), and Collaborative assessments for generating personalized feedback on student learning activities. This framework consists of four key steps. First, peers and professors' assessments are conducted through standardized rubrics (that include both quantitative and qualitative evaluations). Second, multimodal data are collected during learning activities, including video recordings, audio capture, gaze tracking, physiological signals (heart rate, motion data), and behavioral interactions. Third, personalized feedback is generated using AI, synthesizing human-based evaluations and data-based multimodal insights such as posture, speech patterns, stress levels, and cognitive load, among others. Finally, students review their own performance through video recordings and engage in self-assessment and feedback visualization, comparing their own evaluations with peers and professors' assessments, class averages, and AI-generated recommendations. By combining human-based and data-based evaluation techniques, this framework enables more accurate, personalized and actionable feedback. We tested MOSAIC-F in the context of improving oral presentation skills.

1. Introduction

Feedback is a cornerstone of effective learning, serving as a vital tool for students to develop and refine their skills. When delivered thoughtfully, feedback transcends mere correction, guiding students toward deeper understanding, enhanced self-regulation, and sustained progress [1]. Feedback should not be viewed as a one-way transmission of information but as an interactive process that fosters reflection, decision-making, and the development of personal improvement strategies [2].

However, improving students' skills requires structured and effective feedback, which traditional methods often fail to provide adequately [3]. Common challenges include subjectivity in evaluation, where feedback can be inconsistent and overly dependent on personal perspectives rather than on clear, standardised criteria, making it difficult for students to understand exactly what aspects of their performance need improvement. Additionally, a lack of specificity and clarity in feedback comments often results in students receiving vague or generic statements that fail to provide actionable guidance, as they do not explicitly highlight both strengths and weaknesses while offering concrete suggestions on how to enhance their performance.

Another major issue is the absence of opportunities for self-reflection and the application of feedback to future tasks. Many students receive feedback at the end of an assignment or course, with no structured mechanism to integrate it into their learning process. As highlighted in [4], students who reviewed video recordings of their performances became more aware of their strengths and weaknesses, allowing them to identify specific areas for improvement.

In order to address these challenges, we introduce MOSAIC-F, an acronym for a data-driven framework that uses Multimodal Learning Analytics, Observations, Sensors, Artificial Intelligence (AI), and Collaborative assessments to generate personalized feedback on student learning activities. In more detail, in this article, we focus on using MOSAIC-F to improve students' oral presentation skills.

This framework integrates multiple approaches to enhance the assessment process for students' skills combining peer and self-assessment through standardized rubrics that incorporate both quantitative and qualitative evaluations; sensors and Multimodal Learning Analytics (MMLA) [5] to collect data from physiological signals and audio and camera recordings, enabling more accurate and data-based feedback; and AI to address scalability concerns, analyze multimodal data and synthesize evaluations into comprehensive feedback.

To assess the effectiveness of the MOSAIC-F framework, a case study was conducted for enhancing oral presentation skills with final-year students from the Telecommunication Technology and Service Engineering program at Universidad Autónoma de Madrid (UAM). The study focused on evaluating how multimodal data, peer and professor collaborative assessment, and AI-generated feedback could contribute to the development of professional communication competencies in engineering students.

The study involved 46 students who were required to deliver a 10-minute oral presentation followed by a 5-minute question period as part of their course assessment. Participation was voluntary, and all students signed informed consent forms detailing the data collection process, including video/audio recordings and the use of wearable sensors. Furthermore, ethical considerations were strictly observed, and all collected data were fully anonymized at all times to ensure the protection of participants' privacy.

The remainder of the article is organized as follows: Section 2 reviews related work in the areas of Multimodal Learning Analytics and Artificial Intelligence for automated feedback. Section 3 introduces the MOSAIC-F framework and describes its application in a case study focused on enhancing students' oral presentation skills. Section 4 outlines the multimodal data analyses that will be carried out as part of the case study. Finally, Section 5 presents the conclusions and discusses directions for future work.

2. Related Works

2.1. Multimodal Learning Analytics

Multimodal Learning Analytics (MMLA) has emerged as a promising field that seeks to capture, integrate, and analyze diverse data sources to better understand learning processes and improve educational outcomes [5]. Unlike traditional Learning Analytics [6], which often relies solely on log data from learning management systems, MMLA incorporates rich and heterogeneous modalities such as video, audio, physiological signals, gaze data, and behavioral traces.

MMLA has proven to be highly effective in online learning environments, where platforms based on biometrics and behavioral analysis have emerged [7, 8], benefiting from recent advances in machine learning and digital behavior understanding. For example, [9, 10, 11] presents M2LADS, a web-based system that integrates and visualizes multimodal data from MOOC learning sessions. The system collects and synchronizes biometric signals, such as EEG data, heart rate, and visual attention, with behavioral logs and learning performance indicators, offering instructors a dashboard that provides a comprehensive view of learners' cognitive and emotional engagement during the session. In [12], the authors investigate how biometric and behavioral signals can be used to detect distractions related to mobile phone use during online learning sessions, specifically by analyzing head pose deviations captured in the IMPROVE database [13]. Additionally, in [14] focuses on the analysis of visual attention through eye-tracking data to estimate the specific task the learner is performing.

MMLA has also been successfully applied in face-to-face learning environments. For instance, [15] presents a multimodal system embedded in physical computing worktables, which captures students' hand movements, gaze direction, use of programming interfaces, and audio levels to evaluate collaboration and predict the quality of student-generated artifacts. By applying supervised machine learning techniques, the system was able to identify strong performance predictors, such as physical proximity between students and hand motion dynamics.

Similarly, [16] explored the use of MMLA techniques to analyze face-to-face teaching practices based on classroom audio recordings. Their system combines deep learning for speaker diarization and machine learning for practice classification to identify instructional methods such as lectures or group work.

2.2. Artificial Intelligence and Automatic Feedback

Artificial Intelligence has emerged as a key tool for processing and analyzing multimodal educational data, enabling the detection of hidden patterns, the handling of large data volumes, and the delivery of adaptive feedback [17, 18]. More recently, the rise of generative AI has enhanced the capabilities of automated feedback systems—ranging from the generation of written comments to real-time dashboards and alert mechanisms that monitor students' learning progress [19].

In the case of oral presentations, several studies have focused on developing dashboards that use visualizations to provide students with feedback during practice sessions before delivering their presentations. For instance, in [20], the authors introduced a multimodal feedback system that provides real-time analysis of nonverbal communication, including voice volume, posture, gestures, and pauses. The system uses sensor-based technologies such as depth cameras and microphones to capture and interpret students' nonverbal behavior during a presentation. Based on interviews with public speaking experts, the tool identifies a set of effective and ineffective nonverbal practices, offering automated feedback that aligns with commonly accepted teaching methods. Importantly, the study emphasizes that such systems should prioritize raising students' awareness and encouraging reflection, rather than enforcing rigid performance standards.

Building on this line of work, in [21] an AI-driven tool that supports both the content creation and rehearsal phases of oral presentations was developed and it offers guidance on structuring messages, provides audio and video-based feedback on delivery, and uses self-reflection prompts to foster awareness and improvement.

Similarly, [22] proposed a low-cost solution that leverages basic sensors, such as webcams and ambient microphones, to analyze key features like gaze direction, posture, voice volume, filled pauses, and visual slide content. Their system generates post-session feedback reports, integrating video and audio recordings to help students identify areas for improvement.

More recently, [23] presented an open-source multimodal system designed to provide automated feedback on oral presentation skills in real time. The system evaluates body language, voice volume, articulation speed, gaze direction, filled pauses, and visual slide design, offering presenters immediate and actionable information. Unlike many prior tools, this system emphasizes accessibility, modularity, and scalability, aiming to support widespread adoption in educational settings and enable further research into automated feedback mechanisms.

Additionally, virtual reality has also been explored in this context [24, 25]. For example, [25] proposed a VR-based rehearsal system that automatically evaluates students' presentations using machine learning techniques applied to gesture and movement data captured in immersive environments, providing feedback and enabling self-review through avatar playback.

2.3. Evaluating Human Versus AI-Based Feedback

One of the critical aspects in the implementation of automated feedback is how users perceive its credibility, usefulness, and fairness. Prior research comparing human tutors with intelligent tutoring systems has shown that automated feedback can be nearly as effective as human-generated feedback in certain contexts [26]. However, important differences emerge in terms of acceptance: students often perceive human-generated feedback as more trustworthy and contextually sensitive [27]. [28] further supports these concerns, showing that students perceived auto-generated feedback as less valuable and emotionally disconnected, especially when they were aware that it was produced by a language model.

Additionally, in [29] a large-scale study was conducted with over 450 university students to assess how the perceived identity of a feedback provider (human or AI) affects student evaluations. They found

that students often rated AI-generated feedback lower once its source was revealed, especially in terms of genuineness and credibility, even when the content quality was comparable. While students may recognize the benefits of AI, concerns about transparency, fairness, and overreliance can significantly influence their perception, an insight particularly relevant when evaluating how AI-generated feedback is received [30].

Despite these concerns, recent studies suggest that generative AI tools like ChatGPT can produce feedback that is both detailed and effective. In [31], the authors compared human and AI-generated feedback on student essays. While human assessors provided more nuanced and personalized comments, ChatGPT consistently delivered rubric-aligned feedback that met formal assessment criteria with high precision. Similarly, [32] examined students' responses to AI-generated comments on conceptual physics questions. Students rated the AI feedback as accurate and, in many cases, more useful than human-generated feedback due to its comprehensiveness and coverage.

At the same time, generative AI offers clear advantages in terms of scalability and immediacy. Unlike human-generated feedback, which can be time-consuming and inconsistent, AI systems can produce instant draft comments, which instructors can then refine to enhance pedagogical value [33].

Given these developments, recent work emphasizes the need to go beyond system performance and address the human factors surrounding educational technology adoption. In particular, involving educators and learners in the design of AI-based feedback systems has shown promise in enhancing both their usability and acceptance. Co-design approaches facilitate the alignment of data-driven tools with classroom needs, build trust among stakeholders, and promote a stronger sense of ownership and engagement in the use of educational technology [34]. Along these lines, [35] emphasizes pedagogical grounded feedback interventions based on student data, contextual awareness, and personalization, all of which are critical for increasing students' trust and acceptance of automated feedback systems.

3. MOSAIC-F Applied to Enhancing Students' Oral Presentation Skills

In this article, we introduce MOSAIC-F, a data-driven framework that integrates Multimodal Learning Analytics, Observations, Sensors, Artificial Intelligence (AI), and Collaborative assessments for generating personalized Feedback. To illustrate the implementation and potential of MOSAIC-F, we present its application in a concrete case study focused on improving the oral presentation skills of students. In this case study, MOSAIC-F involves both professors and students in the feedback process through the following roles:

- **Evaluators:** At least one professor and two students evaluate the presenter's performance.
- **Presenter:** The student who delivers the oral presentation.
- **Observers:** One student in the audience is monitored using an eye-tracking device, while a research assistant simultaneously annotates key events during the presentation, such as moments of nervous movement, instances of reading from notes or slides, and episodes of eye contact with the audience.

MOSAIC-F use a four step workflow (Figure 1):

1. Peers and Professors' Assessment (see Subsection 3.1)
2. Multimodal Data Collection (see Subsection 3.2)
3. Feedback Generation (see Subsection 3.3)
4. Self-Assessment and Feedback Visualization (see Subsection 3.4)

3.1. Peers and Professors' Assessment

In this step, students' performance is evaluated. In our case study, oral presentations are assessed by both professors and peers using a standardized rubric within the AICoFe system [36]. AICoFe provides individual web-based dashboards for each evaluator, where the rubric is presented. These dashboards are accessible through any web browser.
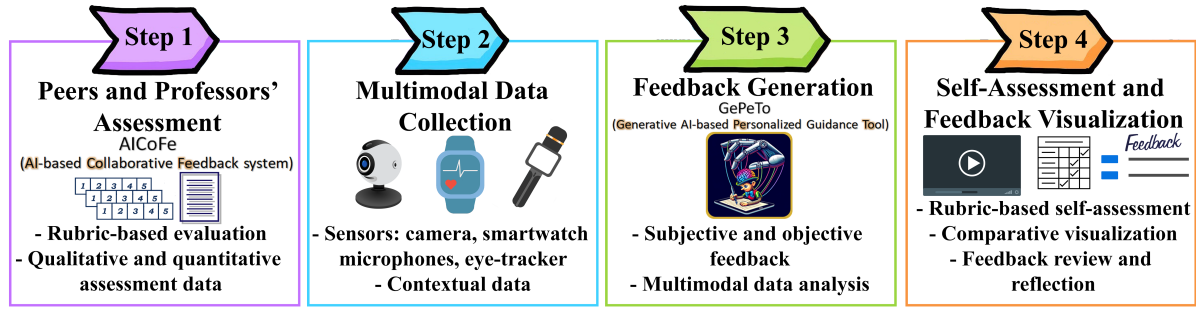
Figure 1: MOSAIC-F workflow diagram.

The rubric includes different items, such as eye contact, attention capture or clarity of the opening, and each item is assessed using a 5-point Likert scale, with predefined descriptions for each level to ensure consistency among all the evaluators. Additionally, peers and professors should provide qualitative observations for each item.

3.2. Multimodal Data Collection

MOSAIC-F uses several sensors and data as illustrated in Figure 2 to monitor students and gain data-based insights into their performance while delivering an oral presentation. In particular, the following sensors are used:

- **Two Logitech C920 PRO HD Webcams:** One RGB webcam records the presenter's performance from a front-facing view. The second webcam records the evaluators and the external observer during the oral presentation. Through the edBB platform [8], both audio and video from these webcams are captured.
- **Integrated Webcams:** Presenters and evaluators are recorded using Microsoft Teams and the webcams integrated in their laptops.
- **Poly SYNC 40 Microphone:** This non-intrusive ambient microphone is placed near the student to capture high-quality audio of the oral presentation using Microsoft Teams.
- **Fitbit Sense Smartwatch:** The presenter, evaluators, and external observer wear a smartwatch that captures heart rate and motion data, including gyroscope, accelerometer, and device orientation.
- **Clicker, Mouse and Keyboard:** Presenters can use a clicker, the mouse or the keyboard to advance the slides. All interactions with these devices are stored.
- **Keyboard and Interactions Logs:** While evaluators assess oral presentations using AICoFe [36], all clicks and keystrokes made on the platform are recorded.
- **Tobii Pro Glasses 3:** An observer wears an eye-tracking device that captures data on gaze, fixations, and saccades.
- **Contextual Data Annotations:** An observer (research assistant) uses a custom interface to label contextual events during the presentation in real time, such as instances of nervous movement, reading behavior, and moments of eye contact. These structured annotations enrich the multimodal dataset and can be later integrated into machine learning models for behavior analysis.

3.3. Feedback Generation

Once all evaluations are collected, AICoFe leverages Generative AI through GePeTo [37] to automatically generate personalized feedback based on the quantitative and qualitative input provided by peers and professors. This feedback is structured around three core components: (1) a summary of the presenter's strengths, (2) identification of areas for improvement, and (3) an action plan that offers concrete, targeted recommendations for enhancing oral presentation skills.
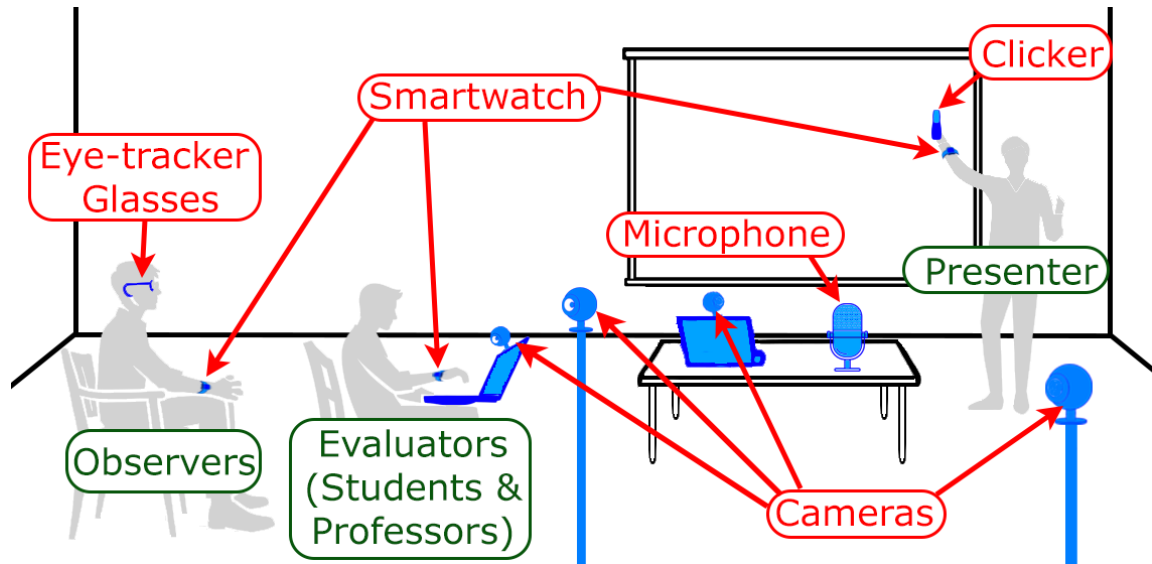
Figure 2: Acquisition setup illustrating the different roles and sensors used.

GePeTo is built on a fine-tuned version of the ChatGPT language model, specifically adapted for oral presentations. The model has been fine-tuned using feedback examples, ensuring that the generated outputs are pedagogically sound, contextually relevant, and aligned with the evaluation rubric used in AICoFe.

In addition to this human-based feedback, a data-based feedback report is generated based on MMLA, incorporating multiple analyses derived from the captured multimodal data. Section 4 presents the analyses to be included in our case study of oral presentations.

3.4. Self-Assessment and Feedback Visualization

A key component of effective feedback is supporting students in reflecting on their own performance. In this step of the MOSAIC-F framework, students review video recordings of their performances as a foundation for self-assessment. This practice allows them to observe themselves from an external perspective, recognize specific behaviors, and evaluate their performance more objectively. Once the self-assessment is submitted, personalized feedback is provided along with visualizations that compare their results to those of peers and professors, as well as class averages. Finally, students are invited to reflect on the feedback received and indicate whether they agree with the assessment and find it useful for their improvement process.

In our case study, after delivering their presentation, students engage in a self-assessment process using video recordings captured from both Microsoft Teams and the frontal camera. These recordings allow them to critically review their own performance and complete the same standardized evaluation rubric used by peers and professors within the AICoFe system. Upon submitting their self-assessment, the system displays the personalized feedback. Although the feedback is automatically produced, it is subsequently reviewed by professors to ensure its accuracy, coherence, and alignment with the rubric. This human oversight helps prevent potential inconsistencies or hallucinations from the language model, ensuring that the final feedback remains pedagogically valid and trustworthy.

In addition to textual feedback, AICoFe provides interactive visualizations that allow students to compare their self-assessment results with peer and professor evaluations, as well as with class averages. This comparative analysis helps students gain deeper insight into their performance, identify discrepancies between self and external assessments, and reflect on areas for growth.

Finally, this reflection process is further supported by a final step in which students are asked to evaluate the feedback they received, indicating whether they found it accurate, useful, and relevant for improving their future presentations.

4. Multimodal Data Analyses in the Oral Presentations Case Study

In Step 3 (see Subsection 3.3), MOSAIC-F generates data-based feedback using multimodal and biometric data. In our case study, we are planning to conduct the following analyses:

- **Head Pose Analysis:** edBB and Microsoft Teams recordings will be used to estimate the Euler angles (pitch, yaw, and roll). The presenter's head pose will then be analyzed to infer patterns of visual attention throughout the presentation. This analysis will help identify when the presenter is making eye contact with the audience, directing their gaze toward the projected slides, looking downward at the floor, or referring to written notes. Similarly, the head pose of the evaluators will also be analyzed to assess their level of attentiveness during the evaluation process. By examining the direction and stability of their gaze, it will be possible to infer when evaluators were actively focused on the presenter, when they were evaluating, and when they were distracted.

- **Body Posture Analysis:** Once the body landmark data have been extracted, the presenter's posture will be analyzed to identify various types of body language and the specific moments at which they occur. For instance, closed or constrained positions—such as crossed arms, hunched shoulders, or limited mobility—may convey a lack of self-confidence and negatively affect the effectiveness of the presentation. In contrast, excessive or erratic movements, including repetitive pacing or continuous shifting of weight, could signal nervousness and potentially distract the audience. On the other hand, open and stable postures will be associated with confidence, clarity, and stronger audience engagement.

- **Audio Analysis:** The presenter's vocal performance will be analyzed. Key audio features, including tone (pitch), modulation, fluency, and clarity, will be extracted to evaluate aspects of speech delivery. This analysis will help identify monotone speech, lack of vocal variation, and articulation issues, all of which are critical for effective oral communication.

- **Speech Transcription and Pattern Detection:** The presenter's speech will be automatically transcribed. This transcription will enable the detection of linguistic patterns that may affect communication quality, including the frequent use of filler words (e.g., "um", "uh", "like", or "you know"), false starts, and long or frequent pauses.

- **Heart Rate Analysis:** The presenter's heart rate will be analyzed using data collected from the smartwatch. The signal will be processed to extract temporal patterns and identify fluctuations throughout the presentation. The analysis will focus on detecting peaks, stable periods, and abrupt changes that may correspond to key segments such as the introduction, transitions, or audience interactions. To evaluate whether heart rate significantly varies across different presentation phases, statistical tests such as paired t-tests will be applied. These tests will allow for the comparison of heart rate levels between predefined intervals (e.g., opening vs. conclusion), helping to determine whether physiological stress responses are consistently associated with specific moments of the presentation. In addition, peaks will be detected and mapped against the presentation timeline to determine whether they coincide with specific events, such as the start of the talk, slide changes, or audience questions. Heart rate data will also be analyzed for the evaluators to assess their physiological engagement during the presentation.

- **Gaze Analysis:** Gaze data from an external observer wearing eye-tracking glasses and located in the audience will be analyzed. The recorded data will include precise information on gaze direction, fixations, eyeblinks and saccades in real-world coordinates. This information will be used to determine whether the presenter successfully captured and maintained the observer's visual attention throughout the presentation. By mapping gaze fixations onto areas of interest (AOIs), such as the presenter's face, the slides, or irrelevant regions, the analysis will reveal how effectively the presenter directs audience attention. Temporal patterns of gaze distribution will also be examined to identify shifts in focus, lapses in attention, or moments of distraction.

- **Keystroke and Interaction Logs:** While evaluators assess presentations through the AICoFe system, all keystrokes and interaction events will be logged with precise timestamps. This interaction data will be used to analyze the sequence and timing of evaluation steps, providing

evidence of whether the assessment process was carried out appropriately. For instance, it will be possible to detect if an evaluator rated final rubric items (such as conclusions) before the corresponding segment of the presentation occurred, indicating a possible lack of attention or premature judgment. Additionally, typing behavior and the time spent on each rubric item will be analyzed to reveal the level of engagement and depth of reflection during the evaluation. By examining these logs, the analysis will support quality assurance in the assessment process, helping to identify superficial, rushed, or inconsistent evaluations.

- **Slides Analysis:** In addition to tracking slide transitions during the presentation, the original `.pptx` file will be analyzed to examine the structure and design of the slides. This analysis will include the extraction of key features such as the total number of slides, slide titles, use of images, presence and position of slide numbers, and the estimated text density per slide. Particular attention will be paid to the font size used in text boxes, as overly small fonts may hinder readability and negatively affect audience engagement.

5. Conclusions and Future Work

In this article, we introduced MOSAIC-F framework, a novel approach to enhancing students' skills through the integration of Multimodal Learning Analytics, Observations, Sensors and Collaborative assessment along with Artificial Intelligence that supports the entire process. Its structured four-step process comprising peer and professor assessments, multimodal data collection, AI-driven feedback generation, and student self-assessment with feedback visualization and aims to deliver more comprehensive, personalized and actionable feedback on student performance.

By combining human-based assessments with multimodal data, such as posture, gaze, speech patterns, and physiological signals, MOSAIC-F has the potential to provide richer, more personalized feedback that goes beyond traditional evaluation methods. The inclusion of self-reflection activities and comparative visualizations further supports student engagement, metacognition, and the development of targeted improvement strategies.

We carried out a case study focused on improving oral presentation skills in a university face-to-face learning setting, applying the MOSAIC-F framework. This initial implementation allowed us to validate the framework's feasibility and laid the groundwork for future iterations. We plan to apply MOSAIC-F to additional case studies across diverse educational contexts and competencies to further evaluate its adaptability, effectiveness, and impact on student learning.

As part of our future work, we will conduct an in-depth analysis of the multimodal data collected during the case study to assess the effectiveness and accuracy of the feedback mechanisms integrated in MOSAIC-F. Beyond this, we aim to expand the framework's application to diverse learning scenarios and to different skill domains such as teamwork. Additionally, we plan to develop interactive tools that allow students to independently practise their skills and receive formative feedback, supporting ongoing improvement beyond formal assessment contexts.

Acknowledgments

Declaration on Generative AI

The authors used ChatGPT-4 in order to grammar and spelling check.

References

[1] J. Hattie, H. Timperley, The Power of Feedback, Review of Educational Research 77 (2007) 81–112.

[2] S. Askew, S. Askew, Feedback for Learning, Technical Report, RoutledgeFalmer London, 2000.

[3] M. Henderson, T. Ryan, M. Phillips, The Challenges of Feedback in Higher Education, Assessment & Evaluation in Higher Education (2019).

[4] M. Tailab, N. Marsh, Use of Self-Assessment of Video Recording to Raise Students' Awareness of Development of Their Oral Presentation Skills, Higher Education Studies 10 (2020).

[5] M. Giannakos, D. Spikol, D. Di Mitri, K. Sharma, X. Ochoa, R. Hammad, The Multimodal Learning Analytics Handbook, Springer, 2022.

[6] C. Lang, G. Siemens, A. Wise, D. Gasevic, Handbook of Learning Analytics (2017).

[7] X. Baró-Solé, A. E. Guerrero-Roldan, J. Prieto-Blázquez, A. Rozeva, O. Marinov, C. Kiennert, P.-O. Rocher, J. Garcia-Alfaro, Integration of an Adaptive Trust-Based E-Assessment System into Virtual Learning Environments—The TeSLA Project Experience, Internet Technology Letters 1 (2018) e56.

[8] R. Daza, A. Morales, R. Tolosana, L. F. Gomez, J. Fierrez, J. Ortega-Garcia, edBB-Demo: Biometrics and Behavior Analysis for Online Educational Platforms, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 37, 2023, pp. 16422–16424.

[9] A. Becerra, R. Daza, R. Cobos, A. Morales, M. Cukurova, J. Fierrez, M2LADS: A System for Generating Multimodal Learning Analytics Dashboards, in: 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE, 2023, pp. 1564–1569.

[10] A. Becerra, R. Daza, R. Cobos, A. Morales, J. Fierrez, M2LADS Demo: A System for Generating Multimodal Learning Analytics Dashboards, arXiv preprint arXiv:2502.15363 (2025).

[11] A. Becerra, R. Daza, R. Cobos, A. Morales, J. Fierrez, User Experience Study Using a System for Generating Multimodal Learning Analytics Dashboards, in: Proceedings of the XXIII International Conference on Human Computer Interaction, 2023, pp. 1–2.

[12] A. Becerra, J. Irigoyen, R. Daza, R. Cobos, A. Morales, J. Fierrez, M. Cukurova, Biometrics and Behavioral Modelling for Detecting Distractions in Online Learning, in: Proc. Simposio Internacional de Informática Educativa (SIIE), VII Congreso Español de Informática, 2024.

[13] R. Daza, A. Becerra, R. Cobos, J. Fierrez, A. Morales, IMPROVE: Impact of Mobile Phones on Remote Online Virtual Education, arXiv preprint arXiv:2412.14195 (2024).

[14] M. Navarro, A. Becerra, R. Daza, R. Cobos, A. Morales, J. Fierrez, VAAD: Visual Attention Analysis Dashboard Applied to E-Learning, in: 2024 International Symposium on Computers in Education (SIIE), IEEE, 2024, pp. 1–6.

[15] D. Spikol, E. Ruffaldi, G. Dabisias, M. Cukurova, Supervised Machine Learning in Multimodal Learning Analytics for Estimating Success in Project-Based Learning, Journal of Computer Assisted Learning 34 (2018) 366–377.

[16] F. P. García, O. Cánovas, F. J. G. Clemente, Exploring AI Techniques for Generalizable Teaching Practice Identification, IEEE Access (2024).

[17] N. Bosch, Y. Chen, S. D'Mello, It's Written on Your Face: Detecting Affective States from Facial Expressions While Learning Computer Programming, in: Intelligent Tutoring Systems: 12th International Conference, ITS 2014, Honolulu, HI, USA, June 5-9, 2014. Proceedings 12, Springer, 2014, pp. 39–44.

[18] C. C. Ekin, O. F. Cantekin, E. Polat, S. Hopcan, Artificial Intelligence in Education: A Text Mining-Based Review of the Past 56 Years, Education and Information Technologies (2025) 1–43.

[19] M. Kim, S. Kim, S. Lee, Y. Yoon, J. Myung, H. Yoo, H. Lim, J. Han, Y. Kim, S.-Y. Ahn, et al., LLM-Driven Learning Analytics Dashboard for Teachers in EFL Writing Education, arXiv preprint arXiv:2410.15025 (2024).

[20] J. Schneider, D. Börner, P. Van Rosmalen, M. Specht, Presentation Trainer: What Experts and Computers Can Tell About Your Nonverbal Communication, Journal of Computer Assisted Learning 33 (2017) 164–177.

[21] D. Di Mitri, J. Schneider, N. Mouhammad, S. Hummel, M. Alomari, M. A. Ali, M. H. R. Masum, H. Arif, M. Rose, R. Klemke, Enhance Your Presentation Skills with Presentable, in: Proceedings of

the 15th International Conference on Learning Analytics & Knowledge (LAK 2025), Demo Track, Dublin, Ireland, 2025. March 3–7.

[22] X. Ochoa, F. Domínguez, B. Guamán, R. Maya, G. Falcones, J. Castells, The RAP System: Automatic Feedback of Oral Presentation Skills Using Multimodal Analysis and Low-Cost Sensors, in: Proceedings of the 8th International Conference on Learning Analytics and Knowledge, 2018, pp. 360–364.

[23] X. Ochoa, H. Zhao, OpenOPAF: An Open-Source Multimodal System for Automated Feedback for Oral Presentations, Journal of Learning Analytics 11 (2024) 224–248.

[24] R. Daza, L. Shengkai, A. Morales, J. Fierrez, K. Nagao, SMARTe-VR: Student Monitoring and Adaptive Response Technology for e-learning in Virtual Reality, in: Proc. AAAI Workshop on Artificial Intelligence for Education, 2025.

[25] Y. Yokoyama, K. Nagao, VR Presentation Training System Using Machine Learning Techniques for Automatic Evaluation, International Journal of Virtual and Augmented Reality (IJVAR) (2021).

[26] K. VanLehn, The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems, Educational Psychologist 46 (2011) 197–221.

[27] H. Zhang, D. Litman, Co-Attention Based Neural Network for Source-Dependent Essay Scoring, arXiv preprint arXiv:1908.01993 (2019).

[28] S. Rüdian, J. Podelo, J. Kužílek, N. Pinkwart, Feedback on Feedback: Student's Perceptions for Feedback from Teachers and Few-Shot LLMs, in: Proceedings of the 15th International Learning Analytics and Knowledge Conference, 2025, pp. 82–92.

[29] T. Nazaretsky, P. Mejia-Domenzain, V. Swamy, J. Frej, T. Käser, AI or Human? Evaluating Student Feedback Perceptions in Higher Education, in: European Conference on Technology Enhanced Learning, Springer, 2024, pp. 284–298.

[30] T. Nazaretsky, P. Mejia-Domenzain, V. Swamy, J. Frej, T. Käser, The Critical Role of Trust in Adopting AI-Powered Educational Technology for Learning: An Instrument for Measuring Student Perceptions, Computers and Education: Artificial Intelligence (2025) 100368.

[31] J. Steiss, T. Tate, S. Graham, J. Cruz, M. Hebert, J. Wang, Y. Moon, W. Tseng, M. Warschauer, C. B. Olson, Comparing the Quality of Human and ChatGPT Feedback of Students' Writing, Learning and Instruction 91 (2024) 101894.

[32] T. Wan, Z. Chen, Exploring Generative AI Assisted Feedback Writing for Students' Written Responses to a Physics Conceptual Question with Prompt Engineering and Few-Shot Learning, Physical Review Physics Education Research 20 (2024) 010152.

[33] C. D. Kloos, C. Alario-Hoyos, I. Estévez-Ayres, P. Callejo-Pinardo, M. A. Hombrados-Herrera, P. J. Muñoz Merino, P. M. Moreno-Marcos, M. Muñoz Organero, M. B. Ibáñez, How Can Generative AI Support Education?, in: 2024 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2024, pp. 1–7.

[34] H. Ogata, C. Liang, Y. Toyokawa, C.-Y. Hsu, K. Nakamura, T. Yamauchi, B. Flanagan, Y. Dai, K. Takami, I. Horikoshi, et al., Co-Designing Data-Driven Educational Technology and Practice: Reflections from the Japanese Context, Technology, Knowledge and Learning 29 (2024) 1711–1732.

[35] P. Topali, A. Ortega-Arranz, Y. Dimitriadis, S. Villagrá-Sobrino, A. Martínez-Monés, J. I. Asensio-Pérez, Unlock the feedback potential: Scaling effective teacher-led interventions in massive educational contexts, in: Innovating Assessment and Feedback Design in Teacher Education, Routledge, 2023, pp. 1–19.

[36] A. Becerra, R. Cobos, Enhancing the Professional Development of Engineering Students Through an AI-Based Collaborative Feedback System, in: 2025 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2025, pp. 1–9.

[37] A. Becerra, Z. Mohseni, J. Sanz, R. Cobos, A Generative AI-Based Personalized Guidance Tool for Enhancing the Feedback to MOOC Learners, in: 2024 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2024, pp. 1–8.