

Review

Not peer-reviewed version

E-Mote: An Emotion-Aware Teacher Training Framework Integrating Facs, AI, and VR

[Rosa Pia D'Acri](#)^{*}, Francesca Bertacchini, [Eleonora Bilotta](#), [Alessandro Soranzo](#)

Posted Date: 15 October 2025

doi: 10.20944/preprints202510.1201.v1

Keywords: teacher training; teacher noticing;; social and emotional learning(SEL); facial action coding system (FACS); artificial intelligence (AI); virtual reality (VR); micro-expressions



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

E-MOTE: An Emotion-Aware Teacher Training Framework Integrating FACS, AI, and VR

Rosa Pia D'Acri ¹, Francesca Bertacchini ², Eleonora Bilotta ³ and Alessandro Soranzo ³

¹ Department of Cultures, Education and Society, University of Calabria, Cosenza, Italy

² Department of Mechanical Engineering, University of Calabria, Cosenza, Italy

³ Department of Physics, University of Calabria, Cosenza, Italy

* Correspondence: rosapia.dacri@unical.it

Abstract

This paper presents E-MOTE (*Emotion-aware Teacher Education Framework*), a conceptual framework designed to enhance teacher education through the integration of the Facial Action Coding System (FACS), Artificial Intelligence (AI), and Virtual Reality (VR). Grounded in neuroscientific and educational research, the proposed framework aims to strengthen teachers' emotional awareness, teacher noticing, and social and emotional learning (SEL) competencies. E-MOTE outlines a design-based approach to addressing persistent gaps in current teacher preparation programs, which often lack tools for practicing real-time recognition of subtle emotional cues. As a theoretical and design-oriented proposal, it provides a structured foundation for developing emotionally responsive teaching and inclusive classroom management, while outlining essential ethical safeguards and scalable validation strategies for diverse educational contexts.

Keywords: teacher training; teacher noticing; social and emotional learning (SEL); facial action coding system (FACS); artificial intelligence (AI); virtual reality (VR); micro-expressions

1. Introduction

Teaching is one of the most emotionally complex and interpersonally demanding professions. Educators operate in environments where subtle emotional cues significantly influence student motivation, attention, and learning outcomes (Damasio, 1994; Hargreaves, 2000; Immordino-Yang, 2016b). The ability to perceive and respond to students' emotional states, especially micro-expressions lasting less than half a second, is a critical competency for classroom management and inclusive pedagogy (Porter & Ten Brinke, 2018).

Despite extensive research on social and emotional learning (SEL) for students, teacher preparation programs still rarely provide systematic training in emotional perception. Most existing approaches emphasize theoretical understanding of emotional intelligence but do not offer structured opportunities to practice real-time recognition and adaptive responses in safe, feedback-rich contexts (Jennings & Greenberg, 2009; Liu & Wang, 2023). This gap is mirrored in educational technology: while AI and VR show promise for teacher training, current systems primarily focus on macro-behaviors and scripted interactions. They lack the granular, validated, and real-time feedback on micro-expressions necessary for training the subtlest aspects of emotional perception (Billingsley et al., 2023a; Chiu et al., 2023b).

Empirical research has consistently shown that teachers often struggle to accurately recognize students' emotional states, which can negatively affect classroom climate and learning outcomes. For instance, Zurbriggen et al. (2023) reported low levels of accuracy in teachers' judgments of students' well-being, while Halberstadt et al. (2022) provided experimental evidence that preservice teachers often commit systematic errors when decoding children's facial expressions.

This paper proposes the E-MOTE (Emotion-aware Teacher Education Framework), a novel conceptual framework that integrates three complementary components to address this specific gap:

- FACS (Facial Action Coding System) for scientifically validated decoding of subtle facial cues;
- Artificial intelligence (AI) for automated analysis and adaptive, real-time feedback;
- Virtual reality (VR) for immersive, controlled environments in which teachers can safely experiment with emotionally complex classroom situations.

Conceptually, the proposed E-MOTE framework builds on three complementary foundations:

1. the ability-based model of emotional intelligence (Mayer et al., 2016);
2. design-based research (DBR) methodologies (McKenney & Reeves, 2012);
3. simulation-based teacher-training environments such as SimSchool (Gibson et al., 2007; Clift & Choi, 2020).

E-MOTE is presented as a research-informed conceptual proposal, bridging interdisciplinary theory and educational practice. It functions as a conceptual model and a protocol for future pilot implementations, remaining open to iterative refinement and empirical validation.

State of the art: AI and VR in teacher training

Although interest in technologies supporting emotion recognition in education is growing, current literature remains fragmented. Most contributions investigate emotion recognition separately from artificial intelligence (AI) and virtual reality (VR), with only a few attempts at systematic integration. Bibliometric and visual analyses confirm this fragmented landscape, showing both rapid growth and persistent silos in research on AI in education.

AI-based emotion recognition has been applied in educational contexts for emotion classification, learning analytics, and adaptive instruction (Luckin et al., 2016; Luckin, 2018). Luckin and colleagues argue that AI should be designed not to replace human intelligence but to augment it, fostering the development of ethical, emotional, and cognitive skills in teachers and learners. In line with this perspective, D’Mello et al. (2017) and Chiu et al. (2023c) highlight AI’s potential to personalize learning based on students’ affective states. However, most systems still function as black boxes, offer limited actionable feedback, and rarely employ validated decoding tools such as FACS.

VR shows strong potential for cultivating teachers’ socio-emotional competencies through *high-fidelity simulations*. Research by Rodríguez-Andrés et al. (2022) demonstrates VR’s effectiveness in simulating complex classroom dynamics that foster empathy, classroom management and reflective capacity. Yet current VR applications rarely incorporate instant emotion recognition or AI-driven feedback, limiting their contribution to training teachers’ perceptual micro-skills for decoding subtle expressions during live instruction.

Recent work has attempted to integrate AI-driven affective feedback into VR-based teacher training. For instance, Zhang et al. (2023) explore hybrid environments where teachers interact with AI-powered avatars capable of emotional responsiveness. While promising, these models generally omit validated decoding systems like FACS and do not address the need for systematic micro-expression analysis. To the best of our knowledge, no *replicable* and *systematically documented* model currently integrates these elements into a coherent, ethically grounded framework. The proposed E-MOTE framework explicitly employs FACS to decode emotional micro-expressions, embeds AI-powered real-time feedback into VR training scenarios, grounds implementation in ethical standards (General Data Protection Regulation, GDPR) and articulates pedagogical pathways for professional development and classroom application.

Table 1 provides a descriptive overview of representative studies on AI and VR in teacher training. The studies selected for comparison were chosen for being both recent (published within the 2022-2024 window, representing the current technological frontier) and representative of the dominant approaches in the literature. The selected contributions illustrate the main technological trends while also revealing persistent gaps.

Studio	FACS decoding	AI real-time feedback	VR immersion	Pedagogical integration	Ethical safeguards
Chiu et al. (2023)	X	✓	X	X	X

Studio	FACS decoding	AI real-time feedback	VR immersion	Pedagogical integration	Ethical safeguards
Rodríguez-Andrés et al. (2022)	X	X	✓	Partial	X
Makransky & Mayer (2022)	X	X	✓	✓	X
Zhang et al. (2023)	X	✓	✓	Partial	X
E-MOTE	✓	✓	✓	✓	✓

Legend: ✓ = present and explicitly implemented; X = absent; Partial = included but not central or lacking methodological clarity.

For the purpose of this analysis, "Pedagogical Integration" is assessed based on the explicit and central connection between the technological tool and a defined pedagogical theory or a structured pathway for professional development. A rating of '✓' (Present) indicates that the study explicitly grounds its application in educational theory (e.g., experiential learning, reflective practice) and outlines how the technology leads to the development of specific, transferable teaching competencies. A rating of 'Partial' indicates that while educational benefits are mentioned, the paper lacks a clear, methodological link to a pedagogical framework or a detailed model for how the tool fosters professional growth.

For example, Chiu et al. (2023a) emphasize AI's potential to personalize learning, though their systems remain focused on student modelling rather than teacher training. Rodríguez-Andrés et al. (2022) demonstrate the effectiveness of VR for simulating classroom dynamics but without AI-driven feedback. Taken together, these studies highlight both the promise of AI-VR integration and the persistent absence of robust and ethically grounded approaches to micro-expression analysis.

E-MOTE is presented as a research-informed conceptual framework bridging interdisciplinary theory and educational practice. It functions both as a conceptual model and as a protocol for future pilot implementations, remaining open to iterative refinement and empirical validation. The rest of this paper is structured to elaborate on this proposal as follows. First, Section 1 establishes FACS as the methodological foundation for decoding micro-expressions. Section 2 then details the integration of AI and VR, outlining the operational workflow (2.1) and advanced use cases (2.2) of the proposed AI-FACS-VR module. Next, Section 3 discusses the core professional competencies E-MOTE aims to cultivate, followed by an analysis of their projected pedagogical impacts in Section 4. Section 5 then examines application prospects and systemic integration, situating E-MOTE against existing models like SimSchool. Finally, the paper concludes by outlining limitations and a future research agenda in Section 6.

2. FACS as a Tool for Decoding Emotional Micro-Expressions in Educational Contexts.

The E-MOTE framework begins with the Facial Action Coding System (FACS), a standardized and validated methodology for objectively decoding the brief, involuntary micro-expressions that reveal authentic emotional states during teacher-student interactions. These micro-expressions, lasting between 1/25 and 1/4 of a second (Ekman & Friesen, 1978; Porter & Ten Brinke, 2018), are critical in educational settings, as students may attempt to conceal feelings of confusion, frustration, or anxiety. FACS enables the identification of specific facial muscle movements, termed Action Units (AUs), which serve as reliable indicators of these underlying emotions. For educators, training in FACS provides a scientific lens to sharpen their perception, moving from a general impression of a student's state to a precise, actionable understanding.

Research has demonstrated that with proper training, observers can significantly improve their ability to detect and interpret these subtle facial cues (Matsumoto & Hwang, 2011). For a teacher, this skill is not about becoming a clinical diagnostician, but about gaining a more accurate and timely understanding of the classroom climate and individual student needs.

Table 2. Pedagogically relevant facial action units (aus) and their implications.

AU code & name	Muscle movement	Inferred emotional state	Potential pedagogical significance & response
AU1: inner brow raiser	Eyebrows raised inward	Worry, sadness, concentration	May indicate a student is struggling with a concept. A teacher could respond with a clarifying question, offer encouragement, or check for understanding
AU4: brow lowerer	Eyebrows lowered/drawn together	Frustration, anger, intense focus	Signals rising frustration or cognitive load. This is a key moment to intervene with scaffolding, a short break, or by re-framing the task to prevent disengagement.
AU5: upper lid raiser	Upper eyelid raised, eye widening	Upper eyelid raised, eye widening	Fear, surprise, vigilance

It is crucial to note that emotions are typically expressed through the co-occurrence of multiple AUs rather than single markers. For instance:

- Sadness often involves AU1, AU4, and AU15 (lip corner depressor), signalling a need for empathetic connection or support.
- Genuine happiness or engagement is reflected in the concurrent activation of AU6 (cheek raiser) and AU12 (lip corner puller), the "Duchenne smile," which can serve as positive feedback for a teacher's instructional approach.
- Boredom or disengagement is often signaled by AU17 (chin raiser) and AU25 (lips part), indicating a need to increase the lesson's pace, interactivity, or relevance.

By focusing on these targeted combinations of AUs, educators can learn to recognize nuanced emotional cues with greater accuracy. This enables them to foster more empathetic and pedagogically responsive interactions, such as providing timely support to a frustrated student before they give up, or recognizing when a quiet student's expression indicates confusion rather than comprehension.

This perceptual skill is both contextual and holistic. It arises from the interplay between local facial cues and global perceptual organization, shaped by neural and sociocultural factors (Adolphs, 2002, 2009; Wagemans et al., 2012). Neuroaesthetic studies, such as those on the expressive ambiguity of Leonardo's portraits, reinforce that meaning is derived from the dynamic interaction of unstable facial cues within a whole (Soranzo & Newberry, 2015, 2016; Soranzo, 2022; 2024). This aligns with the dynamic integration of perceptual and emotional processing in the brain (Zeki, 1999). Furthermore, cross-cultural research shows that identical facial configurations can elicit divergent interpretations, revealing differences in recognition thresholds and affective meanings (Jack et al., 2012b; Chen & Jack, 2017). This variability underscores the critical importance of interpreting facial cues through culturally responsive lenses (Immordino-Yang et al., 2016a; Ting-Toomey, 1999), a principle that is foundational to the adaptable design of the E-MOTE framework.

3. Integrating AI with VR into Teacher Training

Building on the FACS foundation, this section details how E-MOTE leverages AI and VR to create a responsive training ecosystem. The framework moves beyond theoretical knowledge by providing teachers with immersive practice and granular feedback on the perceptual micro-skills of emotion recognition. The integration is designed to bridge the gap between knowing about emotions and expertly perceiving them in the dynamic, fast-paced context of a classroom.

Affective computing, defined by Picard (1997) as the use of computational systems to recognize, interpret, and influence emotions, provides the interdisciplinary backbone for this integration. By leveraging FACS and computer vision, AI can automate facial expression analysis with high accuracy. A significant advance has been the integration of FACS with deep learning architectures, which enables the rapid decoding of facial muscle activations. Within teacher education, these technologies are harnessed not just for analysis, but to promote reflective practice and self-regulated professional growth through targeted feedback.

Recent research highlights the benefits of combining VR with affective computing. Building on these insights, the E-MOTE framework aims to offer a comprehensive, culturally calibrated approach designed for scalable application in both pre-service and in-service teacher training.

3.1. Operational Workflow of the AI-FACS-VR Module

The following workflow outlines the core AI-FACS-VR component of E-MOTE, illustrating how raw facial data from a VR simulation is processed and transformed into actionable pedagogical feedback for the teacher-in-training.

Step 0 – Ethical and privacy safeguards: before any data processing, E-MOTE requires explicit ethical safeguards. This includes informed consent, data minimization, pseudonymization, and bias audits, in compliance with GDPR and United Nations Educational, Scientific and Cultural Organization (UNESCO, 2023) recommendations. Crucially, participants are informed that the system is a training tool for use only within controlled simulations. Building on recent advances in explainable artificial intelligence, E-MOTE ensures algorithmic transparency and interpretability, so that analytical and feedback processes remain understandable and accessible to educators (Adadi & Berrada, 2018).

Step 1 – Data acquisition: within the VR environment, virtual student avatars are equipped with realistic, FACS-compliant facial rigging. The system generates and records the corresponding AU configurations for these avatars based on the pre-programmed emotional responses to the teacher-trainee's actions.

Step 2 – Identification of AUs: computer vision algorithms, specifically Convolutional Neural Networks (CNNs) trained on annotated datasets like CK+ (Lucey et al., 2010), analyze the avatar's facial expressions in real-time to identify the activated AUs.

Step 3 – Processing and classification of emotions: the system maps the identified AUs to inferred emotional states (e.g., AU1+AU4 → sadness/worry; AU6+AU12 → genuine engagement). Deep neural networks (DNNs) enhance the accuracy of interpreting these AU combinations within the specific context of the classroom scenario.

Step 4 – Real-time feedback to the teacher-trainee: this is the critical pedagogical step where data becomes coaching. The AI integration provides feedback to the teacher in two primary, non-intrusive forms:

In-simulation cue: a subtle, color-coded halo or icon appears above the avatar's head (e.g., amber for "frustration" indicated by AU4, blue for "confusion" indicated by AU1). This allows the teacher to practice recognizing cues and adjusting their approach during the interaction.

Post-simulation debriefing dashboard: after the scenario, the teacher reviews a timestamped log of the emotional states they encountered, paired with video clips of their interactions. This dashboard highlights moments where key micro-expressions were present and allows for reflection on their responses, directly fostering the "reflective growth" competency.

This pipeline constitutes the core of E-MOTE's training process, transforming a VR simulation into a reflective learning experience. Empirical evidence supports this pedagogical potential: immersive virtual reality environments have been shown to enhance learners' sense of presence and engagement, fostering deeper emotional and cognitive processing during simulation-based activities (Makransky & Mayer, 2022).

3.2. Advanced Use Cases of the AI-FACS-VR Integration

Building on the operational workflow, this subsection presents advanced use cases, demonstrating how the framework can be applied to complex, authentic teacher education scenarios.

- Practicing inclusive classroom management: VR enables the creation of highly realistic simulations that replicate the complexities of diverse classrooms. These immersive environments foster reflective awareness by allowing teachers to observe their own behavioral responses in emotionally charged situations.
 - Scenarios may include mediating interpersonal conflicts or de-escalating a frustrated student (practicing responses to sustained AU4 and AU7). Teachers can interact with avatars representing students with emotional regulation difficulties, allowing them to practice co-regulation strategies in a safe space.
 - Pedagogical transfer: the immediate feedback on AUs helps teachers calibrate their ability to detect early signs of disengagement (e.g., AU17) or anxiety (AU1+AU20), enabling them to make timely instructional decisions like reframing a task or offering validation *before* a student fully disengages.
- Developing culturally responsive cue recognition: simulations can be designed to model culturally specific nuances in emotional expression, training teachers to avoid misinterpretations. The AI feedback can be calibrated to different cultural datasets, making the teacher aware that the intensity or meaning of an AU configuration (like AU12 for a smile) can vary.
- Formative assessment of teacher noticing: the system provides aggregated, objective data on a teacher's "noticing" skills. For example, a post-session report could reveal that a teacher consistently missed subtle cues of confusion (AU1) in a particular student avatar, indicating a specific area for growth in their perceptual acuity.

In this context, the synergistic use of AI and VR provides an advanced technical-pedagogical solution for simulating and analyzing complex educational scenarios (Holmes et al., 2019). These experiences foster the metacognition and self-regulated learning essential for developing transferable, context-sensitive teaching skills.

4. Core Professional Competencies and Ethical Foundations

Having introduced the technological integration of AI and VR, this section shifts to the professional competencies that E-MOTE is designed to cultivate. Specifically, the framework fosters four core teaching competencies, developed through AI-FACS feedback and VR simulations and framed within ethical and pedagogical principles. They are:

- Emotional attunement: the ability to accurately recognize students' facial cues, including micro-expressions, with the support of AI-based detection systems, and to integrate this information to adapt communicative tone and teaching strategies in real time. This aligns with the ability model of emotional intelligence, which defines EI as a set of cognitive-emotional skills including perception, understanding, and regulation of emotions (Mayer et al., 2008; 2016).
- Empathic responsiveness: practiced through immersive VR scenarios that allow teachers to simulate context-sensitive reactions and emotionally supportive behaviors, a process shown to enhance learning outcomes through emotionally engaging and cognitively structured experiences (Makransky & Petersen, 2021a; Parong & Mayer, 2018), and supported by evidence

on the role of presence and immersion in fostering psychological engagement in VR environments (Slater, 2018).

- Reflective growth: fostered through self-assessment dashboards and post-simulation analytics that translate micro-expression data into personalized development goals, in alignment with a vision of teacher learning as inquiry-based and situated within communities of practice (Cochran-Smith & Lytle, 1999).
- Ethical inclusion: grounded in Nussbaum's (2010) capabilities approach, which emphasizes justice, empathy, and respect for the dignity and potential of every learner.

Trained iteratively through the AI-FACS-VR pipeline, these competencies align with evidence that SEL programs improve classroom climate and outcomes (Durlak et al., 2011). Beyond individual competencies, E-MOTE emphasizes a broader ethical stance toward the design and adoption of educational technologies. Zawacki-Richter et al. (2019) observed that excluding teachers from AI system design creates a risk of pedagogical misalignment. Subsequent literature confirms this risk, noting consequences such as mistrust, poorly aligned tools, and missed educational objectives. Accordingly, the ongoing shift toward human-centred AI in education emphasizes participatory design approaches that prioritize safety, reliability, and trustworthiness (Alfredo et al., 2024).

5. Projected Pedagogical Impacts and Classroom Implications

Building on the core competencies and the AI-FACS-VR workflow outlined previously, this section examines the projected pedagogical impacts of the E-MOTE framework. The central proposition is that the repeated, deliberate practice of identifying specific Action Units (AUs) within immersive VR scenarios will build the mental models and perceptual automaticity necessary for improved teacher responsiveness in real classrooms.

This capacity is projected to emerge from the framework's unique combination of granular feedback and high-fidelity simulation. The AI-FACS-VR module is designed to train educators to detect fleeting non-verbal cues—such as the subtle activation of AU1 (inner brow raiser) indicating worry or AU4 (brow lowerer) signaling rising frustration—that often precede disengagement or conflict. By providing immediate feedback on these precise cues during a simulation, E-MOTE aims to recalibrate a teacher's perceptual system, making them more sensitive to these signals in real time.

The pedagogical impact lies in transforming this heightened awareness into instructional action. For example, recognizing the early signs of confusion (AU1) in a student may guide the teacher to provide a differentiated explanation before the student falls behind. Similarly, detecting the onset of frustration (AU4, AU23) during a challenging task can prompt timely intervention, such as offering encouragement or breaking the task into smaller steps, thus providing emotional scaffolding and sustaining engagement.

These scenarios illustrate the plausible mechanisms through which E-MOTE is designed to strengthen reflective, affect-attuned pedagogy. The framework is theoretically expected to shorten the feedback loop between student expression and teacher response, enhancing interpretative accuracy. Insights from SEL research suggest that this level of emotionally attuned teaching contributes to improved classroom climate and learning outcomes, particularly for vulnerable students (Jennings & Greenberg, 2009; Hargreaves, 2000). More recent evidence confirms that mindfulness-based SEL interventions can enhance teachers' emotional balance and foster positive, supportive classroom environments (Valbusa et al., 2022). These theoretical insights constitute the pedagogical foundation of E-MOTE, reinforcing its design as a teacher education framework that promotes emotionally responsive and inclusive practice.

6. Application Prospects and Systemic Integration

Moving from classroom practice to broader contexts, this section examines the application prospects of E-MOTE and situates it in relation to established approaches, most notably SimSchool.

While E-MOTE shares with SimSchool a foundation in simulation-based training, it proposes a significant evolution in scope and methodology.

As demonstrated in Table 3, E-MOTE builds upon the established theoretical foundations of simulation-based training but extends them by targeting a previously unaddressed layer of teacher competency: the real-time perception of subtle, non-verbal emotional cues. Compared with SimSchool's focus on scripted scenarios and macro-behaviors, E-MOTE's integration of FACS and real-time feedback offers a pathway to train the rapid, subconscious perceptual processes that underpin empathy and attunement.

Table 3. A comparative analysis of SimSchool and the proposed E-MOTE framework.

Dimension	SimSchool	E-MOTE
Primary training focus	Classroom management and instructional decision-making using AI-driven learner profiles	Perceptual acuity and emotional responsiveness through validated micro-expression decoding, immersive VR, and AI-driven adaptive feedback
Underlying approach	Relies on scripted learner behaviors and pre-programmed scenarios, not real affective data	Data-driven emotional analytics using FACS to decode authentic, non-verbal cues in real-time, combined with competency-based training pathways
Feedback mechanism	Primarily post-simulation analysis of teacher actions on student macro-behaviors (e.g., "engagement" score)	Real-time, granular feedback on perceptual performance (e.g., "You missed a cue of frustration (AU4) in Student A"), provided both in-VR and via post-simulation analytics
Technological components	AI-driven behavioral simulation of students	Unified ecosystem: FACS, AI analytics, VR immersion, multimodal feedback, and culturally responsive calibration
Scope of innovation	A mature platform for simulating classroom dynamics and pedagogical strategies	Advances the field by targeting the foundational skill of emotional perception, uniting validated micro-expression analysis, immersive practice, and adaptive feedback in a single training ecosystem designed for scalability

Application prospects for this proposed framework can be articulated across three domains:

- Initial teacher training: universities could integrate E-MOTE modules to provide foundational training in emotional competence, combining theoretical FACS instruction with immersive VR simulations to build perceptual skills from the outset of a teacher's career.
- Ongoing professional development: for in-service teachers, E-MOTE could offer advanced modules focused on specific challenges, such as managing inclusive classrooms or recognizing cross-cultural emotional expressions, using the high-fidelity simulations for deliberate practice.
- Tool for reflective inquiry: the post-simulation dashboards provide objective data on a teacher's "noticing," making it a powerful tool for self-assessment and coaching within professional learning communities.

Looking ahead, realizing these prospects will require careful attention to teacher readiness and cognitive load. Avoiding overload necessitates not only technical refinements but also evidence-based guidelines for integrating such training into existing programs, in line with neuroscience insights on the interplay between emotion, cognition, and learning (Howard-Jones, 2014).

Overall, E-MOTE has the potential to evolve from a conceptual proposal into a systemic innovation in teacher education, provided its implementation remains participatory, ethical, and empirically validated. Its strength resides in its targeted integration of FACS-based micro-expression analysis, AI-driven feedback, and immersive VR within a unified, ethically grounded training model.

A key priority is the active involvement of students in innovation processes. Engaging learners as stakeholders in the design and evaluation of educational interventions not only enhances effectiveness but also fosters a more inclusive and participatory culture. Systematically integrating student voice into educational practices reinforces the pedagogical validity by ensuring responsiveness to students' experiences and evolving needs (Cook-Sather, 2006).

7. Limitations and Future Research Agenda

As a conceptual framework, the primary limitation of E-MOTE is that it awaits empirical validation. While grounded in established research from neuroscience, education, and affective computing, its efficacy and feasibility must be tested through rigorous implementation. Prior research on emotionally attuned teaching has shown clear benefits (Hargreaves, 2000), but these findings need to be empirically extended to the specific AI-FACS-VR training context that E-MOTE proposes.

To guide this necessary transition from concept to evidence-based tool, a multi-phase research agenda is outlined below, designed to address the core challenges:

Phase 1: technical validation and feasibility

- Objective: to assess the accuracy and reliability of the AI-FACS-VR module in a controlled laboratory setting and establish its basic usability.
- Research priorities:
 1. Algorithmic accuracy: evaluate the performance of the FACS-decoding AI when applied to the facial animations of virtual student avatars, ensuring it can accurately identify pedagogically relevant AUs in real-time.
 2. User experience (UX) and usability: conduct studies with pre-service teachers to assess the usability of the VR interface and the perceived usefulness of the feedback mechanisms (e.g., in-VR cues, post-simulation dashboard) using think-aloud protocols and standardized UX scales.
 3. Mitigating bias: initiate the development of culturally diverse, FACS-annotated datasets of virtual expressions to audit and mitigate algorithmic bias, ensuring the system does not misclassify emotions based on avatar ethnicity or gender (Mehrabi et al., 2021).

Phase 2: pedagogical validation and short-term efficacy

- Objective: to evaluate the framework's impact on teacher learning and competency development in controlled training contexts.
- Research priorities:

1. Impact on noticing skills: employ experimental designs with pre- and post-assessments to determine if training with E-MOTE leads to significant improvements in teachers' accuracy and speed in detecting micro-expressions in standardized video clips of students.
2. Perceived utility and ethical concerns: explore teachers' perceptions of the tool's pedagogical integration, ethical implications, and alignment with professional values through mixed-methods studies (surveys, focus groups).
3. Cross-cultural calibration: initiate cross-cultural validation by testing the framework across diverse educational contexts, examining how cultural norms influence the interpretation of AUs and adapting feedback mechanisms accordingly, in line with the *OECD PISA Global Competence Framework* (Organisation for Economic Co-operation and Development, 2019).

Phase 3: ecological and longitudinal validation

- Objective: to assess the transfer of trained skills to real classrooms and the long-term impact on teaching practice and student outcomes.
- Research priorities:
 1. Transfer to practice: conduct longitudinal studies that track teachers from VR training into their classrooms, using observational methods to see if improved perceptual acuity in simulation leads to more responsive and inclusive teaching behaviors in practice.
 2. Impact on classroom climate and student outcomes: investigate the downstream effects of E-MOTE training by measuring its correlation with improved classroom climate, student engagement, and student self-reports of well-being.

Scalability and implementation science: study the systemic requirements for large-scale adoption, including cost-effectiveness, institutional buy-in, and the development of sustainable trainer-of-trainer models. These ethical commitments are consistent with the Ethical Guidelines on the Use of Artificial Intelligence and Data in Teaching and Learning for Educators issued by the European Commission (2022).

Addressing core challenges across phases

This research must run in parallel with ongoing efforts to address the critical limitations identified:

- Data privacy and governance: all phases must implement and refine the proposed ethical safeguards (granular consent, data minimization, pseudonymization) and subject them to independent ethical review.
- Teacher readiness and technostress: research must investigate the optimal support structures, such as structured professional development and technical coaching, needed to foster pedagogical fluency and reduce technostress.
- Health and safety: best practices for VR use (e.g., session limits, structured debriefings) must be empirically tested and integrated into the protocol to minimize simulator sickness and visual fatigue (Rebenitsch & Owen, 2016).

Funding and partnerships for this ambitious agenda could be sought through European educational innovation programs such as Erasmus+ and Horizon Europe, in collaboration with teacher education institutions, AI research laboratories, and VR developers. Ultimately, the success of E-MOTE depends on this rigorous, iterative research process, ensuring that technological innovation is firmly anchored in pedagogical need and empirical evidence.

8. Conclusions

This paper has introduced E-MOTE, a comprehensive conceptual framework designed to enhance teacher education by systematically targeting the often-overlooked competency of emotional perception. Grounded in the ability-based model of emotional intelligence (Mayer et al., 2016) and design-based research methodologies (McKenney & Reeves, 2012), E-MOTE proposes a novel integration of three core components: the Facial Action Coding System (FACS) for validated micro-

expression decoding, Artificial Intelligence (AI) for real-time analytics and adaptive feedback, and Virtual Reality (VR) for immersive, safe practice environments.

The framework's principal contribution lies in its structured approach to bridging a critical gap between theoretical knowledge of emotions and the embodied skill of perceiving them in the dynamic flow of classroom interaction. By moving beyond the macro-behavioral focus of existing simulation platforms like SimSchool, E-MOTE outlines a pathway to train the perceptual micro-skills—such as recognizing the subtle signs of confusion (AU1) or frustration (AU4)—that underpin emotional attunement and empathic responsiveness.

As a conceptual proposal, E-MOTE provides a structured foundation for developing emotionally responsive teaching practices while explicitly embedding essential ethical safeguards from its inception. It functions as both a theoretical model and a practical protocol for future development, outlining a clear, multi-phase research agenda to guide its transition from concept to validated tool. This agenda prioritizes technical validation, pedagogical efficacy, and ecological transfer, ensuring that future iterations are grounded in empirical evidence.

Ultimately, E-MOTE outlines a promising and ethically informed pathway for advancing teacher education. By uniting validated emotional decoding, immersive simulation, and adaptive feedback within a single training ecosystem, the framework holds significant potential to strengthen the socio-emotional core of teaching. Future work must now focus on the empirical validation, cross-cultural calibration, and strategic implementation of this model, aligning it with broader European digital education and professional development frameworks to explore its potential for global adaptation.

References

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160
- Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 21–62
- Adolphs, R. (2009). The social brain: neural basis of social behavior. *Annual Review of Psychology*, 60, 693–716
- Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). *Human-Centred Learning Analytics and AI in Education: a Systematic Literature Review*. *Computers & Education Artificial Intelligence*, 6(5), Article 100215
- Azukas, M. E., & Kluk, D. (2022). Simulated teaching: an exploration of virtual classroom simulation for preservice teachers during the COVID-19 pandemic. In C. R. Graham, D. Krutka, & S. Kimmons (Eds.), *Exploring online learning through simulations and gaming* (pp. 97–112). Springer
- Badiee, F., & Kaufman, D. (2015). Design evaluation of a simulation for teacher education. *SAGE Open*, 5(2), 2158244015592454
- Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: hidden biases of good people*. Delacorte Press
- Billingsley, G., Smith, S., Smith, S., & Meritt, J. (2023a). A systematic review of immersive virtual reality in teacher education: current applications and future directions. *Journal of Research on Technology in Education*, 55(1), 106–128
- Billingsley, B., Bertram, C., & Nassaji, M. (2023b). How should AI be taught in schools? Ethical and pedagogical issues to consider. *Frontiers in Education*, 8, Article 1145665
- Blakemore, S. J., & Frith, U. (2005). *The learning brain: lessons for education*. Blackwell Publishing
- Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, 22(2), 218–224
- Chen, C., & Jack, R. E. (2017). Discovering cultural differences through emotion perception, learning, and reverse correlation. *Current Opinion in Psychology*, 17, 44–48
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: a meta-analysis. *Review of Educational Research*, 90(4), 499–541

- Chiu, T. K. F., Lin, T.-J., & Lonka, K. (2023a). A systematic review of research on artificial intelligence applications in higher education: learning effectiveness and future directions. *Computers and Education: Artificial Intelligence*, 4, 100142
- Chiu, T. K. F., Hew, K. F., & Ng, C. S. L. (2023b). A systematic review of artificial intelligence applications in K–12 education: Learning outcomes, learner characteristics, and pedagogical strategies. *British Journal of Educational Technology*, 54(2), 357–378
- Chiu, T. K. F., Lin, T.-J., Chai, C. S., Pak, R., & Zhan, Y. (2023c). Artificial intelligence in education: a systematic review and future research directions. *Computers and Education: Artificial Intelligence*, 4, 100108
- Christensen, R., Knezek, G., Tyler-Wood, T., & Gibson, D. (2011). SimSchool: an online dynamic simulator for enhancing teacher preparation. *Journal of Technology and Teacher Education*, 19(3), 277–292
- Clift, R. T., & Choi, S. (2020). *The use of simulation in teacher education: A review of SimSchool and beyond*. *Teaching and Teacher Education*, 91, 103037
- Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: teacher learning in communities. *Review of Research in Education*, 24(1), 249–305
- Cohn, J. F., & Ekman, P. (2005). Observer based measurement of facial expression with the Facial Action Coding System. In J. A. Coan & J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 111–134). Oxford University Press
- Cook-Sather, A. (2006). Sound, presence, and power: "student voice" in educational research and reform. *Curriculum Inquiry*, 36(4), 359–390
- Damasio, A. R. (1994). Descartes' error: emotion, reason, and the human brain. G. P. Putnam's Sons
- Deale, D. F., & Pastore, R. S. (2014). Evaluation of SimSchool: an instructional simulation for preservice teachers. *Journal of Educational Technology Systems*, 42(3), 255–268
- Dimitropoulos, K., Manitsaris, S., & Tsalakanidou, F. (2021). Capturing and analyzing affective features for teacher training on classroom management. *IEEE Transactions on Affective Computing*, 12(3), 790–802
- D'Mello, S., & Graesser, A. (2012). Autotutor and affective autotutor: learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM transactions on interactive intelligent systems (TiiS)*, 2(4), 1–39
- D'Mello, S., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys*, 47(3), 1–36
- D'Mello, S. K., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, and automated: Data science and the future of learning assessment. *Journal of Educational Psychology*, 109(7), 1010–1025
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of schoolbased universal interventions. *Child Development*, 82(1), 405–432
- Ekman, P., & Friesen, W. V. (1972). Cultural differences in facial expression of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation (Vol. 19, pp. 207–283)*. University of Nebraska Press
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200
- Ekman, P. (2003). *Emotions revealed: recognizing faces and feelings to improve communication and emotional life*. Times Books
- European Parliament and Council of the European Union (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1–88
- European Commission. (2022). *Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators*. Publications Office of the European Union
- Gibson, D., Aldrich, C., & Prensky, M. (Eds.). (2007). *Games and simulations in online learning: Research and development frameworks*. Hershey, PA: IGI Global
- Gibson, D. (2011). *SimSchool: An online dynamic simulator for enhancing teacher preparation*. *International Journal of Learning Technology*, 6(2), 201–220
- Gross, J. J. (2015). Emotion regulation: current status and future prospects. *Psychological Inquiry*, 26(1), 1–26
- Halberstadt, A. G., Cooke, A. N., Garner, P. W., Hughes, S. A., Oertwig, D., & Neupert, S. D. (2022). Racialized emotion recognition accuracy and anger bias of children's faces. *Emotion*, 22(3), 403–417

- Hargreaves, A. (2000). The emotional practice of teaching. *Teaching and Teacher Education*, 16(8), 811–826
- Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in education: promises and implications for teaching and learning (1st ed.). Center for Curriculum Redesign
- Hopper, S. B. (2018). Developing teacher know-how through play in SimSchool. *International Journal of Teaching and Learning in Higher Education*, 30(1), 46–56
- Howard-Jones, P. A. (2014). Neuroscience and education: myths and messages. *Nature Reviews Neuroscience*, 15(12), 817–824
- Huang, Y., Li, H., & Fong, R. (2021). The application of artificial intelligence in virtual reality learning environments: a systematic review. *Interactive Learning Environments*, 29(6), 1038–1057
- Immordino-Yang, M. H., Yang, X.F., & Damasio, H. (2016a). Cultural modes of expressing emotions influence how emotions are experienced. *Emotion*, 16(7), 1033–1039
- Immordino-Yang, M. H. (2016b). Emotion, sociality, and the brain's default mode network: insights for educational practice and policy. *Policy insights from the behavioral and brain sciences*, 3(2), 211–219
- Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012a). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, 109(19), 7241–7244
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012b). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: general*, 141(1), 19–25
- Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, 79(1), 491–525
- Kröger, J. L., Meißner, F., & Rannenber, K. (2019). Granular privacy policies with privacy scorecards: transparency of privacy practices through privacy icons. In 2019 IEEE Security and Privacy Workshops (SPW) pp. 62–69
- Liu, Y., & Wang, H. (2023). Teachers' emotional perception and regulation in classroom interactions: Implications for socio-emotional teacher education. *Teaching and Teacher Education*, 125, 104021
- Livingstone, M. (2002). *Vision and art: the biology of seeing*. Harry N. Abrams
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2010, 94–101
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L.B.R. (2016). *Intelligence unleashed: an argument for AI in education*. Pearson Education
- Luckin, R. (2018). Machine learning and human intelligence: the future of education for the 21st century. UCL Press
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: new ability or eclectic traits? *American Psychologist*, 63(6), 503–517
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2016). The ability model of emotional intelligence: principles and updates. *Emotion Review*, 8(4), 290–300
- Makransky, G., & Petersen, G. B. (2021a). The cognitive affective model of immersive learning (CAMIL): a theoretical framework for learning in virtual reality. *Educational Psychology Review*, 33, 937–958
- Makransky, G., & Petersen, G. B. (2021b). Investigating the process of learning with simulation-based virtual reality: a structural equation modeling approach. *Computers & Education*, 166, 104154
- Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. *Educational Psychology Review*, 34, 1771–1798
- Matsumoto, D., & Ekman, P. (1989). American-japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, 13(2), 143–157
- Matsumoto, D., & Hwang, H. C. (2006). Culture and nonverbal behavior. In V. Manusov & M. L. Patterson (Eds.), *The Sage handbook of nonverbal communication* (pp. 219–236). Sage Publications
- Matsumoto, D., & Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motivation and Emotion*, 35(2), 181–191
- McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. Routledge
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), Article 1–35

- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdil, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Buchmann, E., Monreale, A., Pensa, R. G., Dragoni, M., & Hitzler, P. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, 10(3), e1356
- Nussbaum, M. C. (2010). *Creating capabilities: the human development approach*. Harvard University Press
- Organisation for Economic Co-operation and Development. (2019). *OECD PISA global competence framework*. OECD Publishing.
- Pantic, M., & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1424–1445
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality: effects of prior knowledge and learning activity. *Journal of Educational Psychology*, 110(6), 785–797
- Picard, R. W. (1997). *Affective Computing*. MIT Press
- Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: challenges and proposed solutions. *Journal of Cybersecurity*, 4(1), ty001
- Porter, S., & Ten Brinke, L., (2018). Secrets and lies: involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior*, 42(1), 35–56
- Rebenitsch, L., & Owen, C., (2016). Review on cybersickness in applications and visual displays. *Virtual Reality*, 20, 101–125
- Rodriguez-Andrés, D., Juan, M. C., Mollá, R., & Méndez-López, M. (2022). Virtual reality systems as tools for teacher training on emotional competence: a systematic review. *Education and Information Technologies*, 27(4), 5053–5082
- Siegel, D. J. (2020). *The developing mind: how relationships and the brain interact to shape who we are* (3rd ed.). Guilford Press
- Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109(3), 431–433
- Soranzo, A., & Newberry, M. (2015). The uncatchable smile in Leonardo da Vinci's *La Bella Principessa* portrait. *Vision Research*, 113, 78–86
- Soranzo, A., & Newberry, M. (2016). Investigating the “Uncatchable Smile” in Leonardo da Vinci's *La Bella Principessa*: a comparison with the *Mona Lisa* and Pollaiuolo's *Portrait of a Girl*. *Journal of Visualized Experiments (JoVE)*, (116), e54248
- Soranzo, A. (2022). Another ambiguous expression by Leonardo da Vinci. *Gestalt Theory*, 44 (1-2), 41–60
- Soranzo, A. (2024). The psychology of *Mona Lisa's* smile. *Scientific Reports*, 14(1), 12250
- Sutton, R. E., & Wheatley, K. F. (2003). Teachers' emotions and teaching: a review of the literature and directions for future research. *Educational Psychology Review*, 15(4), 327–358
- Tian, Y., Kanade, T., & Cohn, J. F. (2011). Facial expression recognition. In S. Z. Li & A. K. Jain (Eds.). *Handbook of face recognition* (pp. 487–519). Springer
- Ting-Toomey, S. (1999). *Communicating across cultures*. Guilford Press
- United Nations Educational, Scientific and Cultural Organization. (2023). *Recommendation on the ethics of artificial intelligence*. UNESCO Publishing
- Urhahne, D., & Zhu, M. (2015). Accuracy of teachers' judgments of students' subjective well-being. *Learning and Individual Differences*, 43, 226–232
- Valbusa, F., Tagliabue, L., & Vergani, L. (2022). Fostering social-emotional learning in higher education: The impact of a mindfulness-based training program on teacher's well-being and classroom climate. *Frontiers in Psychology*, 13, 928341
- Venetz, M., Zurbruggen, C., & Schwab, S. (2019). What do teachers think about their students' inclusion? Consistency of students' self-reports and teacher ratings. *Frontiers in Psychology*, 10, 1637
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & Von Der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138(6), 1172–1217
- Zhang, Z., Fort, J. M., & Giménez-Mateu, L. (2023). Facial expression recognition in virtual reality environments: Challenges and opportunities. *Frontiers in Psychology*, 14, 1280136

- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in higher education*, 16(1), Article 39
- Zeki, S. (1999). *Inner vision: an exploration of art and the brain*. Oxford University Press
- Zurbriggen, C. L. A., Nusser, L., Krischler, M., & Schmitt, M. (2023). *Teachers' judgment accuracy of students' subjective well-being in school: In search of explanatory factors*. *Teaching and Teacher Education*, 122, 104304

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.