How does the deployment of emotion AI to surveil students in classrooms restructure the governance of educational assessment by shifting decision-making power from pedagogical evidence to algorithmic inference?

June 7, 2026 | SnugLab Research | readme.snuglab.com

Executive Summary

The deployment of emotion AI in classrooms restructures educational assessment governance by shifting decision-making power from nuanced pedagogical evidence to algorithmic inference, primarily through continuous biometric and behavioral data collection. While proponents suggest this shift offers actionable insights and efficiency gains, evidence indicates it introduces significant risks of data distortion, algorithmic bias, and the erosion of teacher agency, thereby degrading the quality and fairness of learning evaluations. The lack of transparency and empirical validation for these systems, coupled with documented disproportionate impacts on minority and neurodivergent students, suggests that this restructuring concentrates authority in opaque automated systems and software vendors rather than enhancing pedagogical practice.

Key Findings

Shift from Pedagogical Evidence to Algorithmic Inference

The deployment of emotion AI fundamentally restructures educational assessment governance by moving decision-making power from traditional pedagogical evidence to algorithmic inference, a process termed "synthetic governance" [2, 3, 13]. Pedagogical evidence relies on human validation, where educators interpret nuanced student behaviors, self-reported feelings, and classroom dynamics to inform instructional decisions [11, 12, 14, 15]. In contrast, algorithmic inference uses Algorithmic Decision Systems (ADS) to analyze vast amounts of personal data, multimodal sensors (e.g., facial expressions, physiological indicators), and pattern recognition to deduce correlations about student emotions, engagement, and cognitive processes [1, 6, 8, 11, 13]. This continuous data stream, often operating invisibly, automates instructional decisions and increases responsibilities delegated to computer systems [2, 3, 6, 7, 13].

Impact on Assessment Validity and Bias

Treating algorithmic inference as functionally equivalent to pedagogical evidence undermines the validity of learning evaluations. Algorithmic inferences often lack empirical validation for their accuracy and effectiveness [4, 5, 7]. These systems are prone to algorithmic bias, particularly misidentifying younger students and people of color, leading to unfair outcomes and systematic discrimination in learning opportunities [2, 4, 5, 6]. Digital monitoring disproportionately affects lower-income, minority, and LGBTQ+ students [4]. For example, facial recognition systems have documented accuracy issues for younger individuals and people of color [5]. While emotion AI can assist neurodivergent learners by decoding nonverbal cues, the continuous surveillance simultaneously creates a "chilling effect" where six in ten students report discomfort in expressing their true thoughts and feelings when monitored [4, 6]. This self-censorship fundamentally alters the emotional data being collected, generating assessment noise and data distortion [4, 6].

Centralization of Authority and Erosion of Teacher Agency

The continuous collection of student biometric and behavioral data establishes a causal chain that displaces teacher judgment, centralizing assessment authority in software vendors and school administrators [2, 3, 7, 13]. Smart classrooms continuously collect data through sensors monitoring facial expressions, eye gaze, body language, and physiological indicators [1, 8]. This data feeds into ADS that infer correlations and automate instructional decisions, often with partial or completely absent human intervention [6, 7]. This data-driven modulation introduces "synthetic governance," shaping learner subjectivities and influencing institutional policy implementation [2, 3, 6, 13]. For instance, algorithms can flag online activity and directly contact police before parents or school administrators are aware, subjecting low-income students to unnecessary policing [4]. This diminishes teacher agency, replacing human intelligence with detached content moderation [4, 13, 14]. However, researchers emphasize that AI should augment, not replace, human intelligence, with teachers' professional judgment remaining central [12, 13, 14, 15].

Actionable Insights vs. Data Distortion

Early emotion AI deployments demonstrate both potential for actionable insights and significant data distortions. Proponents argue that emotion AI provides distinct, actionable

insights by analyzing multimodal data to track student engagement and cognitive processes in real time [1, 8, 11]. These systems can improve teacher efficiency by automating tasks like parent communication, saving educators an estimated 2 to 3 hours weekly [9]. Real-time monitoring of attention patterns allows teachers to dynamically adjust instruction, and behavior-tracking algorithms can identify patterns weeks in advance for preventive support [9]. Studies indicate that students receiving emotionally adaptive feedback perform significantly better in learning tasks [8]. For example, a 2025 study using a Vision Transformer model achieved 90.62% accuracy in recognizing emotional engagement among 40 undergraduate students, linking AI-inferred engagement to quiz-based learning outcomes [18]. Another 2025 study found a convolutional neural network system classified seven emotions with 85% accuracy in university classes, correlating AI-detected happiness with academic performance (r = 0.65) and fear with lower performance (r = âˆ’0.54) [19].

However, these deployments simultaneously generate assessment noise and data distortion. The continuous surveillance creates a "chilling effect," with six in ten students reporting discomfort in expressing their true thoughts and feelings when monitored [4, 6]. This distortion is compounded by cultural misclassification and algorithmic bias, as facial recognition technologies are less accurate for people of color and disproportionately impact lower-income, minority, and LGBTQ+ students [2, 4, 5]. Critics contend there is little empirical proof that these tools are effective or accurate [4, 5, 7].

Transparency and Accountability

The opaque decision-making process of emotion AI erodes transparent assessment governance by preventing students and parents from meaningfully auditing or challenging evaluation outcomes. Emotion AI systems function as "black boxes" that infer correlations without providing clear explanations for individual determinations [2, 3, 6, 13]. This "synthetic governance" operates invisibly within daily educational workflows [2, 3, 7, 13]. True transparency would require access to ADS code and documentation, which is typically incomprehensible to non-experts [6]. This opacity means algorithmic inferences can dictate outcomes without prior notification, such as monitoring algorithms contacting police before parents are aware [4]. To prevent this erosion, experts recommend Algorithmic Impact Assessments (AIA) and maintaining a "human in the loop" to validate inferences against pedagogical evidence [6, 13].

Vendor Landscape and Implementation

Specific emotion AI vendors and deployments demonstrate this shift. Intel and Classroom Technologies' "Class" system automatically analyzes student faces to alert teachers to boredom or confusion [17]. Imentiv's platform analyzes facial expressions, audio tones, and text to detect emotions and modify content difficulty or flag anxiety during exams [8]. A federated learning-based emotion detection system at a university generates aggregated emotion reports for teachers to adjust pedagogical strategies [16]. While specific pricing models and numerical accuracy benchmarks by demographic group are limited, the North Brunswick school district spent over $58,000 in one year on AI monitoring software, and the Lockport City School District invested $3.3 million in a facial recognition system [4, 5]. These systems are known to have accuracy issues and biases against people of color and younger individuals [5].

Implications

The restructuring of educational assessment governance by emotion AI has profound implications for equity, privacy, and the role of human educators. The shift toward algorithmic inference risks embedding and amplifying existing societal biases within educational systems, potentially leading to "algorithmic discrimination" and systematic unfairness in learning opportunities for minority and neurodivergent students [2, 4, 5, 6, 13]. The opacity of these "black-box" systems undermines accountability, making it difficult for students, parents, and even educators to understand or challenge assessment outcomes [2, 6, 7, 10]. Furthermore, the continuous surveillance inherent in emotion AI deployments can create a "chilling effect," altering student behavior and expressions, which in turn corrupts the very data these systems are designed to collect [4, 6, 8]. This erosion of trust and agency, coupled with the high costs of implementation, suggests a need for rigorous ethical frameworks, transparent auditing, and a clear commitment to maintaining human oversight to ensure that AI serves to augment, rather than diminish, the quality and fairness of education [6, 12, 13, 14, 15].

Limitations and Caveats

The available research, while extensive, presents certain limitations. Specific numerical bias rates for neurodivergent students and non-native English speakers in emotion AI systems are not consistently quantified across the findings, nor are direct statistical comparisons to traditional pedagogical assessments in these areas. Furthermore, detailed case studies or pilot programs from 2023-2026 that explicitly measure how

emotion AI directly altered grading outcomes or resource allocation, including the precise percentage of decisions made by algorithms versus teachers, are largely absent. The confidence in the overall impact of emotion AI on educational assessment governance remains moderate, as the topic is empirically testable but significant methodological and ethical debates persist regarding the reliability and impact of these technologies.

Sources

[1] [peer-reviewed] Smart Classrooms: How Sensors and AI Are Shaping Educational Paradigms - Authors: Xiaochen Zhang; Yiran Ding; Xiaoyu Huang; Wujing Li; Liumei Long; Shiyao Ding - Journal: Sensors (Basel, Switzerland) - https://pmc.ncbi.nlm.nih.gov/articles/PMC11397895/

[2] [edu] Section - manifold.umn.edu - https://manifold.umn.edu/read/algorithms-of-education/section/c1774bf0-08e5-4ee7-ac17-876b5b84825e

[3] [edu] WilliamsonB Governing Software And Algorithmic Power In Educ - storre.stir.ac.uk - https://www.storre.stir.ac.uk/bitstream/1893/20498/1/WilliamsonB_Governing%20software%20and%20algorithmic%20power%20in%20education_2014.pdf

[4] Students We Face Invasive Ai Powered School Surveillance Now - aclu-nj.org - https://www.aclu-nj.org/news/students-we-face-invasive-ai-powered-school-surveillance-now-were-calling-lawmakers-regulate-it/

[5] School Surveillance The Students Rights Implications Of Art - northcarolinalawreview.org - https://northcarolinalawreview.org/wp-content/uploads/sites/5/2020/03/School-Surveillance_-The-Students-Rights-Implications-of-Artific.pdf

[6] [gov] EPRS STU(2019)624261 EN - europarl.europa.eu - https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf

[7] View - journal.abcollab.id - https://journal.abcollab.id/index.php/AILLCE/article/view/31

[8] [blog] Emotions In Learning What Ai And Psychology Tell Us About Th - imentiv.ai - https://imentiv.ai/blog/emotions-in-learning-what-ai-and-psychology-tell-us-about-the-classroom/

[9] [blog] Ai Classroom Management Strategies - schoolai.com - https://schoolai.com/blog/ai-classroom-management-strategies

[10] [blog] Empowering Educational Institutions Through Ai Surveillance - security101.com - https://www.security101.com/blog/empowering-educational-institutions-through-ai-surveillance

[11] [blog] Affective Computing The Future Of Education D24d063c0693 - medium.com - https://medium.com/@harveystoner10/affective-computing-the-future-of-education-d24d063c0693

[12] [peer-reviewed] E-MOTE: A Conceptual Framework for Emotion-Aware Teacher Training Integrating FACS, AI and VR - Authors: Rosa Pia D'Acri; Francesco Demarco; Alessandro Soranzo - Journal: Vision - https://pmc.ncbi.nlm.nih.gov/articles/PMC12921945/

[13] [gov] Ai Report - ed.gov - https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf

[14] How Might Ai Transform The Way We Assess And Understand Lear - edtechhub.org - https://edtechhub.org/2025/10/29/how-might-ai-transform-the-way-we-assess-and-understand-learning/

[15] Ai And Formative Assessment - circls.org - https://circls.org/educatorcircls/ai-and-formative-assessment

[16] [peer-reviewed] Implementing federated learning for privacy-preserving emotion detection in educational environments - Authors: GutiÃ©rrez, Rommel; Villegas-Ch, William; LujÃ¡n-Mora, Sergio - Journal: Frontiers in Artificial Intelligence - https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1644844/full

[17] Psychologist Trained Emotion Detection Software - privacyinternational.org - http://privacyinternational.org/examples/4969/psychologist-trained-emotion-detection-software

[18] [peer-reviewed] Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review - AUTHORS UNAVAILABLE - Journal: Frontiers in psychology -

https://pubmed.ncbi.nlm.nih.gov/38966729/
[19] Archivepdf - oajaiml.com - https://www.oajaiml.com/uploads/archivepdf/827742122.pdf