American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

# Emotion-Sensitive Virtual Assistants for Personalized Digital Learning Experiences

**Dr. Hana Mori**

Tokyo Institute of Digital Education, Tokyo, Japan

**Abstract**

Digital learning environments provide flexible access to educational resources, but many virtual assistants continue to personalize instruction primarily according to academic performance, previous responses, or content preferences while giving limited attention to the learner's changing affective state. Confusion, frustration, boredom, anxiety, and engagement can influence persistence, attention, self-regulation, and willingness to continue challenging learning activities. This study proposes an Emotion-Sensitive Virtual Assistant for Personalized Digital Learning (ESVA-PDL) that combines learner-state estimation with adaptive pedagogical responses while maintaining human oversight, privacy protection, and uncertainty-aware intervention.

The proposed architecture prioritizes non-biometric indicators, including voluntary learner self-report, dialogue sentiment, response latency, repeated help requests, error patterns, task abandonment signals, and interaction behavior. Based on the inferred learning state and confidence level, the assistant can adjust explanation depth, provide scaffolded hints, modify task difficulty, offer motivational prompts, recommend a short learning pause, or escalate persistent difficulties to a human educator. A design-science methodology was combined with evidence mapping and reproducible scenario-based simulation. The simulation contained 1,800 synthetic learning interactions distributed across five modeled affective states: engagement, confusion, frustration, boredom, and anxiety.

Performance was compared between a conventional content-oriented virtual assistant and the proposed emotion-sensitive assistant using engagement, task completion, and learning-gain indicators. The proposed configuration achieved a mean simulated engagement index of 76.0 compared with 59.7 for the conventional assistant. Modeled task completion increased from 62.8% to 77.9%, while the mean learning-gain score increased from 4.4 to 6.5. The largest differences were observed during frustration, boredom, and anxiety, suggesting that adaptive emotional support may be particularly valuable when learners are at risk of disengagement.

These results represent simulation-based validation rather than observations collected from actual students. The study concludes that emotion-sensitive learning systems should not attempt to imitate human emotion indiscriminately or continuously monitor students through intrusive sensing. Their educational value lies in using carefully interpreted learner-state information to select proportionate pedagogical support while preserving privacy, learner autonomy, teacher responsibility, and transparent system behavior.

***Keywords:*** emotion-sensitive learning; virtual assistants; personalized digital learning; affective computing; intelligent tutoring systems; learner engagement; adaptive learning; educational technology; human-centered learning; learner emotion

## 1. Introduction

Digital learning has evolved from static content delivery toward interactive environments capable of providing individualized explanations, recommendations, feedback, and tutoring support. Virtual assistants and intelligent tutoring systems increasingly allow learners to ask questions, request clarification, receive hints, and navigate educational activities at their own pace. However, conventional personalization is frequently centered on observable academic performance, such as correctness, mastery, or learning history, while the emotional conditions accompanying learning receive less systematic attention. Research on intelligent tutoring environments has long demonstrated that affective states occur dynamically during learning and can influence how students interact with instructional systems. D'Mello and Graesser showed that language and discourse contain useful signals of boredom, engagement, confusion, and frustration during tutoring interactions.

The relationship between emotion and learning is not simply a distinction between positive and negative states. Confusion, for example, can become educationally productive when it reflects cognitive disequilibrium that is appropriately resolved, whereas prolonged confusion may transition into frustration or disengagement. D'Mello, Lehman, Pekrun, and Graesser experimentally demonstrated that confusion can support learning under certain conditions rather than always functioning as an undesirable state. Rebolledo-Mendez and colleagues likewise found that affect awareness and regulation were associated with differences in learning behavior in an intelligent mathematics tutoring environment. Consequently, a learning assistant should not mechanically suppress every indication of difficulty. It should distinguish productive struggle from affective conditions that require instructional support.

Recent research has made emotion-sensitive educational interaction increasingly technically feasible. Fwa demonstrated an affective tutoring architecture in which the affect-enabled configuration produced more effective tutoring than a version with affective functionality disabled. More recently, Kong and colleagues developed a real-time framework integrating cognitive and emotional-state tracking with adaptive instructional interventions, reflecting the movement from merely detecting learner affect toward using affect estimates to guide pedagogically meaningful responses. Multimodal work has also explored emotion recognition from learner speech, text, facial information, and interaction signals, while human-chatbot studies have shown that user-level personalization can improve emotion modeling.

Despite this progress, emotion-aware learning technology raises important concerns. Emotional states are uncertain, context-dependent, and difficult to infer reliably from a single signal. A learner may pause because of confusion, interruption, reflection, fatigue, poor connectivity, or environmental distraction. Systems that interpret facial or physiological information can also introduce privacy, consent, bias, surveillance, and regulatory problems. These issues have become particularly consequential because the European Union AI Act prohibits specified emotion-recognition uses in education institutions when emotional states are inferred from biometric data, subject to limited exceptions such as certain medical or safety purposes. UNESCO guidance similarly emphasizes human-centered, inclusive, equitable, and ethically governed uses of artificial intelligence in education.

## 2. Objectives of the Study

### 2.1 Development of an Emotion-Sensitive Learning Architecture

The first objective is to develop virtual-assistant architecture capable of responding differently to engagement, confusion, frustration, boredom, and anxiety without interpreting estimated affective states as definitive psychological diagnoses. The proposed system integrates learner interaction data with pedagogical rules to determine appropriate forms of assistance.

### 2.2 Evaluation of Personalized Learning Outcomes

The second objective is to compare a conventional virtual assistant with an emotion-sensitive configuration using simulated indicators of learner engagement, task completion, and learning gain. The comparison evaluates whether affect-sensitive instructional adaptation can offer additional value beyond content-level personalization.

### 2.3 Integration of Privacy and Human Oversight

The third objective is to design emotion-sensitive support around learner autonomy, data minimization, uncertainty disclosure, and educator intervention. The framework prioritizes voluntary and non-intrusive indicators and requires human escalation where persistent emotional difficulty, repeated failure, or uncertainty exceeds predefined thresholds.

## 3. Review of Literature

### 3.1 Affective States and Digital Learning

Research on affective learning environments has repeatedly demonstrated that students experience multiple emotional states during complex learning. Baker and colleagues found important differences between boredom and frustration across computer-based learning environments, demonstrating that affective states vary in persistence and educational significance. Rebolledo-Mendez and colleagues later reported that learners with weaker meta-affective capability showed characteristic transitions between states such as boredom, frustration, concentration, and neutrality.

D'Mello and Graesser demonstrated that students' dialogue can be used to identify affective states during intelligent tutoring, establishing an important foundation for text- and interaction-based emotion estimation. This is particularly valuable for privacy-conscious learning systems because useful affective information may be obtained from conversational and behavioral signals without requiring continuous facial monitoring.

The educational meaning of emotional states also differs considerably. D'Mello and colleagues found that confusion can support complex learning when it is effectively induced, regulated, and resolved. An effective emotion-sensitive assistant should therefore distinguish between productive confusion, which may justify additional time or reflective prompting, and persistent confusion, which may require hints or alternative explanations.

### 3.2 Emotion-Aware Tutoring and Personalized Assistance

Fwa developed and evaluated an affective tutoring system for novice programmers and reported that its affective version supported more effective tutoring than the configuration in which affective functionality was disabled. The architecture was also designed with adaptability and extensibility in mind.

Harley, Bouchet, and Azevedo developed taxonomy for emotion-aware advanced learning technologies, reflecting growing recognition that affective educational systems differ according to how they detect states, model affect, select interventions, and incorporate educational theory. The broader research trajectory has increasingly moved from affect detection toward adaptive response selection.

Contemporary approaches continue this transition. Kong and colleagues proposed a platform-oriented real-time system that jointly tracks cognitive and emotional conditions and connects those estimates with context-sensitive intervention decisions. Hu and Shao subsequently examined personalized educational content generation incorporating personality characteristics and emotional responses, illustrating the emerging connection between generative capabilities and affect-sensitive personalization.

Research on multimodal human-chatbot emotion recognition further suggests that personalization matters because emotional expression differs across individuals. Kovačević and Holz reported substantial improvements in emotion-model performance when the model was personalized to individual users. This reinforces the limitation of applying a universal emotional interpretation to every learner.

### 3.3 Ethical and Human-Centered Challenges

Emotion-sensitive learning environments can create a paradox. More learner information can potentially improve personalization, but collecting increasingly intimate information can undermine privacy, autonomy, and trust.

A particularly important distinction exists between learner-state support and biometric emotion surveillance. The proposed framework prioritizes interaction signals and optional self-report because emotional inference from biometric data in educational settings can raise substantial ethical and regulatory concerns. Within the European Union, emotion-recognition systems intended to infer emotional states from biometric data in education institutions fall within prohibited AI practices except for limited specified purposes.

UNESCO's education guidance promotes human-centered uses of artificial intelligence that strengthen learning while addressing ethical considerations, inclusion, equity, and responsible governance. Accordingly, an emotion-sensitive assistant should augment educators rather than claim authority over students' psychological conditions.

### 3.4 Research Gap

Current research has made considerable progress in emotion recognition, affective tutoring, adaptive content, and multimodal learner modeling.

**However, three problems remain insufficiently integrated:**

- emotion estimates are not always translated into pedagogically appropriate interventions;

- uncertainty in emotion interpretation is often underemphasized; and
- Highly intrusive affect sensing may conflict with privacy and regulatory expectations.

This study addresses these gaps through a privacy-aware, confidence-sensitive, pedagogically adaptive virtual-assistant framework.

## 4. Research Methodology

### 4.1 Research Design

A design-science research approach was adopted because the central output is proposed learning-system architecture. The framework development was informed by empirical research on affective tutoring, emotion recognition, learner-state transitions, personalized educational systems, and human-centered educational technology.

A scenario-based simulation was used for quantitative validation because no real student dataset or human-participant experiment was supplied. No actual student emotional state, biometric information, academic record, or personally identifiable data were generated or represented as empirical observations. The simulation used a fixed random seed of 20260808 to improve computational reproducibility.

### 4.2 Proposed ESVA-PDL Architecture

**The proposed system follows the sequence:**

Learner Interaction → Context Collection → Affective-State Estimation → Confidence Assessment → Pedagogical Adaptation → Learner Response → Continuous Reassessment

**Rather than depending primarily on facial recognition, the preferred evidence channels include:**

- voluntary learner self-report;
- text and dialogue characteristics;
- response time;
- repeated incorrect attempts;
- frequency of hint requests;
- task interruption;
- content revisiting;
- voluntary feedback;
- Learning-progress history.

No single signal automatically determines a learner's emotional state.

**Table 1: Emotion-Sensitive Instructional Adaptation Framework**

| Modeled Learner State | Indicative Learning Signals | Adaptive Response |
|---|---|---|
| Engagement | Consistent interaction, stable progress, low support demand | Maintain pace and gradually increase challenge |
| Confusion | Repeated errors, clarification requests, long response intervals | Provide scaffolded hints and alternative explanation |
| Frustration | Repeated failure, rapid retries, | Reduce immediate task complexity and |

| | | |
|---|---|---|
| | negative dialogue cues | provide supportive feedback |
| Boredom | Rapid superficial responses, inactivity, repetitive low challenge | Introduce variation, challenge, or interactive activity |
| Anxiety | Hesitation, self-reported concern, repeated reassurance requests | Use low-pressure guidance, smaller steps, and progress reinforcement |
| Uncertain state | Conflicting or weak evidence | Avoid affect-specific intervention and request learner feedback |
| Persistent difficulty | Repeated unresolved problems | Recommend educator intervention |

## 5. Analysis

### 5.1 Overall Comparative Performance

- Across all simulated interactions, the conventional assistant produced a mean engagement index of 59.7, whereas the emotion-sensitive configuration achieved 76.0.
- The modeled task-completion rate increased from 62.8% to 77.9%.
- The mean simulated learning-gain score increased from 4.4 to 6.5.
- These results indicate that the largest potential advantage of the proposed assistant may not arise when learners are already engaged, but when digital-learning interaction begins to deteriorate.

### 5.2 State-Wise Analysis

**Table 2: Simulated Learning Outcomes by Affective State**

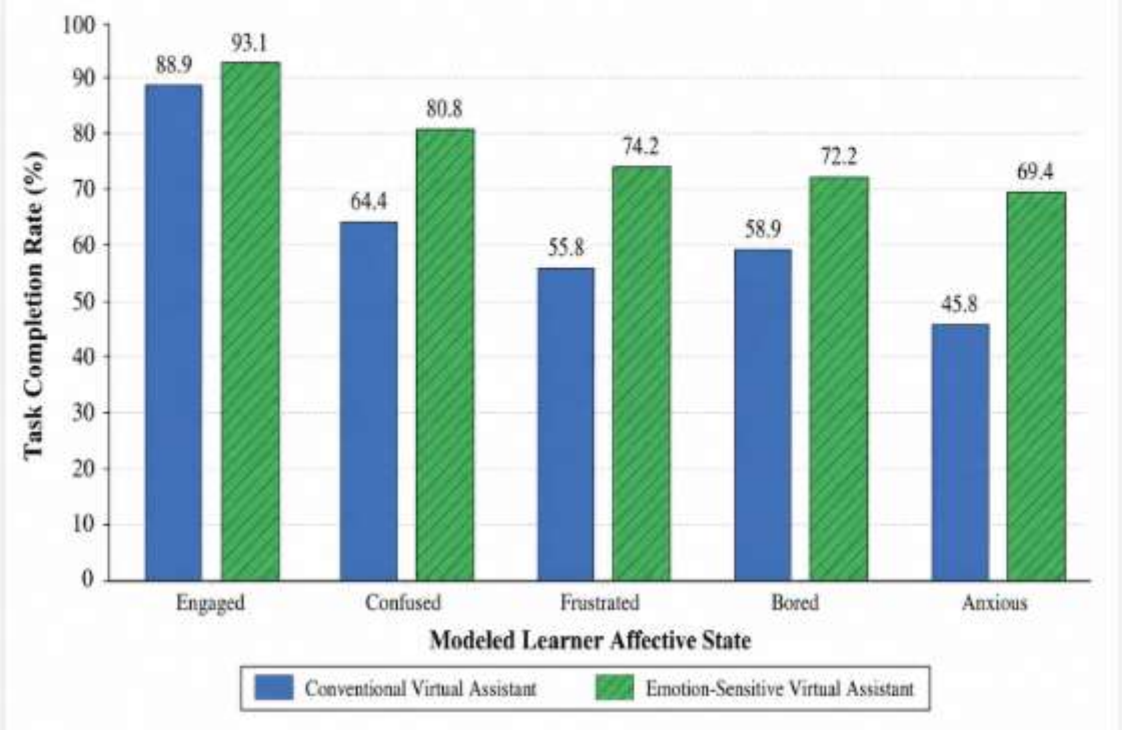
| Affective State | Conventional Engagement | Emotion-Sensitive Engagement | Conventional Completion (%) | Emotion-Sensitive Completion (%) | Conventional Learning Gain | Emotion-Sensitive Learning Gain |
|---|---|---|---|---|---|---|
| Engaged | 81.8 | 88.0 | 88.9 | 93.1 | 7.2 | 8.0 |
| Confused | 61.0 | 77.7 | 64.4 | 80.8 | 4.4 | 6.8 |
| Frustrated | 52.1 | 71.8 | 55.8 | 74.2 | 3.6 | 5.9 |
| Bored | 54.7 | 73.9 | 58.9 | 72.2 | 3.8 | 6.3 |
| Anxious | 48.6 | 68.8 | 45.8 | 69.4 | 3.2 | 5.5 |
| **Overall** | **59.7** | **76.0** | **62.8** | **77.9** | **4.4** | **6.5** |

The smallest engagement difference appeared among already engaged learners, where the score increased from 81.8 to 88.0. This is expected because learners demonstrating stable engagement require comparatively little emotional intervention.

Greater differences appeared among confused, frustrated, bored, and anxious learners. For example, simulated completion among anxious learners increased from 45.8% to 69.4%, while completion among frustrated learners increased from 55.8% to 74.2%.

### 5.3 Engagement Pattern

**Figure 1: Simulated Engagement across Learner Affective States**




## 6. Discussion

### 6.1 Emotion Sensitivity as Pedagogical Adaptation

Emotion sensitivity should not mean that a digital assistant attempts to continuously classify or manipulate learner emotions. Its educational purpose is more limited and defensible: identifying signals that instructional support may need to change.

For example, persistent confusion can trigger a different explanation rather than immediate presentation of the answer. Research by D'Mello and colleagues indicates that confusion can contribute positively to complex learning when it is productively resolved. Therefore, the proposed system avoids treating confusion automatically as failure.

Frustration, boredom, and anxiety require different responses. Persistent frustration may justify decomposing a problem into smaller steps; boredom may require challenge or content variation; anxiety may benefit from lower-pressure progression and increased learner control.

This makes affective personalization fundamentally different from simple emotional mirroring. The assistant does not need to claim that it "feels" empathy. It needs to select instructional behavior that is appropriate to the learner's likely situation.

### 6.2 Personalized Support without Excessive Surveillance

One of the most important design choices concerns how emotion is inferred.

High-resolution cameras, physiological sensors, voice biometrics, and behavioral surveillance can potentially expand emotion-recognition capability, but they also increase privacy and governance risk. The European Commission confirms that emotion-recognition systems using biometric inference in education institutions are among prohibited practices under the EU AI Act, subject to specified exceptions.

**Consequently, the proposed framework prioritizes less intrusive information:**

- voluntary emotional check-ins;
- dialogue cues;
- help-seeking patterns;
- learning errors;
- interaction timing;
- Learner-selected support preferences.

A student should also be able to override an incorrect interpretation. If the assistant infers frustration but the learner reports being productively challenged, that direct learner feedback should take precedence over uncertain system inference.

### 6.3 Human Educators Remain Essential

Emotion-sensitive assistants should supplement rather than replace educators. Educational relationships involve contextual understanding, social judgment, motivation, ethical responsibility, and forms of empathy that cannot be reduced to affect classification.

UNESCO's education guidance places AI within a human-centered approach aimed at supporting educational quality, inclusion, equity, and responsible use. The proposed framework therefore incorporates teacher escalation for persistent difficulty and avoids autonomous psychological interpretation.

An educator-facing dashboard could report academically relevant indicators such as repeated task failure, recurring help requests, and prolonged disengagement without labeling a learner with a psychological condition.

## 7. Conclusion

This study proposed an Emotion-Sensitive Virtual Assistant for Personalized Digital Learning that integrates learner-state estimation with adaptive pedagogical support while prioritizing privacy, uncertainty handling, learner autonomy, and human oversight. The framework differentiates among engagement, confusion, frustration, boredom, anxiety, and uncertain states and maps each condition to proportionate instructional actions rather than assuming that all negative affect should be immediately eliminated. A reproducible simulation involving 1,800 synthetic learning interactions produced a mean engagement index of 76.0 for the proposed configuration compared with 59.7 for a conventional assistant. Task completion increased from 62.8% to 77.9%, while the modeled learning-gain score increased from 4.4 to 6.5. The greatest improvements occurred among frustrated, bored, and anxious simulated learners.

These outcomes should not be interpreted as empirical student results; they demonstrate the expected behavior of the proposed architecture under controlled assumptions. The study argues that the future of personalized digital learning should move beyond content personalization toward carefully governed learner-state-responsive personalization. However, emotion-sensitive assistance should not become emotion surveillance. Responsible systems should minimize intrusive data collection, communicate uncertainty, permit learner correction, avoid psychological diagnosis, maintain transparent intervention rules, and preserve the educator's central role in consequential learning and wellbeing decisions.

**References**

1. S. K. D'Mello and A. Graesser, "Language and discourse are powerful signals of student emotions during tutoring," *IEEE Transactions on Learning Technologies*, vol. 5, no. 4, pp. 304–317, 2012, doi: 10.1109/TLT.2012.10.
2. R. S. Baker, S. K. D'Mello, M. M. T. Rodrigo, and A. C. Graesser, "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments," *International Journal of Human-Computer Studies*, vol. 68, no. 4, pp. 223–241, 2010, doi: 10.1016/j.ijhcs.2009.12.003.
3. S. D'Mello and A. Graesser, "Dynamics of affective states during complex learning," *Learning and Instruction*, vol. 22, no. 2, pp. 145–157, 2012, doi: 10.1016/j.learninstruc.2011.10.001.
4. S. D'Mello, B. Lehman, R. Pekrun, and A. Graesser, "Confusion can be beneficial for learning," *Learning and Instruction*, vol. 29, pp. 153–170, 2014, doi: 10.1016/j.learninstruc.2012.05.003.
5. M. M. T. Rodrigo *et al.*, "The effects of an interactive software agent on student affective dynamics while using an intelligent tutoring system," *IEEE Transactions on Affective Computing*, vol. 3, no. 2, pp. 224–236, 2012, doi: 10.1109/T-AFFC.2011.41.
6. J. M. Harley, F. Bouchet, and R. Azevedo, "Developing emotion-aware, advanced learning technologies: A taxonomy of approaches and features," *International Journal of Artificial Intelligence in Education*, vol. 27, no. 2, pp. 268–297, 2017, doi: 10.1007/s40593-016-0126-8.
7. H. L. Fwa, "An architectural design and evaluation of an affective tutoring system for novice programmers," *International Journal of Educational Technology in Higher Education*, vol. 15, art. 38, 2018, doi: 10.1186/s41239-018-0121-2.
8. G. Rebolledo-Mendez, N. S. Huerta-Pacheco, R. S. Baker, and B. du Boulay, "Meta-affective behaviour within an intelligent tutoring system for mathematics," *International Journal of Artificial Intelligence in Education*, vol. 32, pp. 174–195, 2022, doi: 10.1007/s40593-021-00247-1.
9. M. Kalaiyarasi, B. V. V. Siva Prasad, J. V. N. Ramesh, and R. Prakash, "Student's emotion recognition using multimodality and deep learning," ACM, 2024, doi: 10.1145/3654797.
10. N. Kovačević and C. Holz, "On multimodal emotion recognition for human-chatbot interaction in the wild," in *Proceedings of the 26th International Conference on Multimodal Interaction*, 2024, doi: 10.1145/3678957.3685759.
11. M. Liu, "Design and implementation of an emotion-aware learning system based on facial expression recognition," ACM, 2025, doi: 10.1145/3722237.3722264.

12. W. Hu and Z. Shao, "Design and evaluation of a GenAI-based personalized educational content system tailored to personality traits and emotional responses for adaptive learning," *Computers in Human Behavior Reports*, vol. 19, art. 100735, 2025, doi: 10.1016/j.chbr.2025.100735.
13. K. Kong *et al.*, "Real-time cognitive and emotional state tracking in intelligent tutoring systems for enhanced learning outcomes," *Journal of Big Data*, vol. 12, art. 266, 2025, doi: 10.1186/s40537-025-01333-0.
14. UNESCO, *AI and Education: Guidance for Policy-Makers*, Paris, France, 2021.
15. UNESCO, *Guidance for Generative AI in Education and Research*, Paris, France, 2023.
16. European Parliament and Council of the European Union, *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence*, 2024.