EQīLevel: Emotion-Aware Reinforcement Learning for Adaptive Academic Tutoring

Verónica Elze
School of Technology and Computing
City University of Seattle
Seattle, WA, USA
0009-0006-8417-7744

Taejin Kim
School of Technology and Computing
City University of Seattle
Seattle, WA, USA
0009-0002-2923-6570

Brian Maeng
School of Technology and Computing
City University of Seattle
Seattle, WA, USA
0009-0001-1953-9159

Abstract—Intelligent tutoring systems (ITS) used in cybersecurity education often lack the ability to respond to learners' emotional states during complex analytical tasks. This paper introduces EQīLevel, an emotionally adaptive AI tutoring architecture that integrates reinforcement learning (RL), sentiment detection, and large language model (LLM) dialogue within a lightweight command-line interface (CLI) + FastAPI framework. Traditional intelligent tutoring systems often rely on rigid rule-based structures and rarely account for learners' emotional states, which can reduce engagement and persistence. EQīLevel addresses this limitation by analyzing voice-based cues and dynamically adapting lesson difficulty, tone, and pacing through a JSON-based Model Context Protocol (MCP). The MCP encodes emotion, performance, and learning-style variables into structured state representations that guide Generative Pre-trained Transformer (GPT) dialogue generation and reinforcement learning policy updates. Evaluation using simulated learner interactions demonstrated 78% successful adaptation to frustration scenarios, Whisper transcription accuracy with a 5.3% word error rate (WER), emotion detection accuracy of 84% with 81% tone alignment, and improved reinforcement learning convergence, with average rewards rising from 0.41 to 0.63. In cybersecurity education, EQīLevel illustrates how emotionally adaptive tutoring may help learners remain resilient when confronting ambiguous and adversarial scenarios such as phishing detection and threat analysis. By combining emotional awareness with adaptive instructional control, EQīLevel demonstrates a scalable framework for emotionally adaptive tutoring.

Keywords—*reinforcement learning, emotion-aware tutoring, cybersecurity education, adaptive learning, sentiment analysis, intelligent tutoring systems*

I. INTRODUCTION

This project addresses the limitations of traditional Intelligent Tutoring Systems (ITS) in adapting to learners' emotions, performance, and preferred learning styles in real time. These challenges are amplified in cybersecurity education, where learners must navigate ambiguous threat scenarios and evolving adversarial tactics. EQīLevel is designed as an emotionally adaptive AI tutoring system that integrates RL, real-time sentiment detection, and LLMs within a CLI + FastAPI backend framework. Central to its architecture is the JSON-based MCP, which encodes state variables and controls adaptive dialogue generation. The following subsections preview the problem being addressed, the motivation for the system, the approach taken, and the expected outcomes.

A. Problem Statement

Although ITSs have advanced in personalizing learning experiences, most remain constrained by static content branching and rigid instructional rules. They often fail to incorporate emotional responsiveness or multimodal adaptability, leading to reduced learner engagement, particularly among underserved or self-directed students [1]. Without the ability to detect and respond to changes in a learner's emotional and cognitive state, these systems risk delivering feedback that is poorly timed, mismatched in tone, or ineffective in sustaining motivation. In cybersecurity training, such rigidity can reduce persistence when learners encounter ambiguous log data, incomplete threat indicators, or adversarial deception tactics.

B. Motivation

Research in emotion-aware pedagogy shows that alignment between a learner's emotional state and instructional strategy improves engagement, persistence, and comprehension [2]; [3]. For learners in self-paced or resource-limited environments, emotionally intelligent interaction can act as a proxy for the support a human tutor might provide. EQīLevel is motivated by the need for a scalable, technically robust tutoring system that combines sentiment detection, adaptive instructional control, and modular integration of state-of-the-art AI services to deliver contextually relevant, learner-specific support. In cybersecurity, this adaptability helps sustain engagement during high-stakes exercises where frustration and ambiguity are common, such as phishing detection or incident response simulations.

C. Approach

EQīLevel employs a CLI + FastAPI backend pipeline that captures learner speech, transcribes it via Whisper [4], analyzes sentiment and intent, and encodes these signals into the MCP. The MCP provides an interpretable learner-state representation that stores variables such as emotional tone, task performance, instructional pacing, and difficulty level. These structured variables guide both the reinforcement-learning policy updates and the parameters used by the LLM to generate instructional responses. This architecture allows for dynamic modulation of tone, pacing, difficulty, and instructional style based on real-time emotional and performance data. The approach builds on foundational research in RL-driven tutoring systems [5], interpersonal tutoring agents [6], and emotion-aware instructional agents [3], but extends these capabilities by integrating direct emotional state variables into the policy loop through MCP. This paper illustrates this pipeline in a cybersecurity education use case, where learners practice detecting adversarial behavior under uncertain conditions. Voice and text inputs were selected to keep the prototype lightweight and avoid multimodal sensing requiring human-subject approval.

D. Contributions

This paper makes the following contributions to cybersecurity education research. First, it presents EQīLevel, a unified tutoring architecture that combines emotion detection, reinforcement learning, and LLM-driven dialogue in a single adaptive instructional loop. Second, it introduces a JSON-based Model Context Protocol (MCP) representation that encodes learner emotion, performance, and learning-style signals into an interpretable control state for reinforcement-learning policy adaptation. Third, it situates this architecture within cybersecurity education, a domain characterized by ambiguity, frustration, and adversarial reasoning, where emotionally adaptive instruction may be especially valuable. Fourth, it provides a preliminary simulation-based evaluation showing promising results in real-time adaptation, speech transcription reliability, emotional tone alignment, and reinforcement learning stability.

II. BACKGROUND

Modern ITSs increasingly rely on a combination of machine learning, affective computing, and language technologies to deliver personalized educational experiences. The subsections below define the foundational technologies used in recent educational AI systems: RL, Emotion-Aware AI, Multimodal Learning, LLM-driven instruction, and JSON-based logic control. These descriptions establish the technical baseline for understanding more advanced implementations.

A. Reinforcement Learning (RL)

Reinforcement Learning (RL) enables agents to optimize sequential decisions through trial-and-error interactions guided by reward signals [5]. In educational systems, RL supports dynamic adjustment of instructional pacing and content sequencing. In cybersecurity contexts, RL is especially useful for training learners to respond to ambiguous and evolving threats.

B. Emotion-Aware AI

Emotion-aware AI refers to systems capable of detecting or inferring human affective states through speech, text, facial expressions, or behavioral signals [1]. These signals can be used to adapt instructional responses in educational systems to improve learner engagement and persistence. In cybersecurity education, recognizing frustration during complex labs or incident simulations is critical to sustaining motivation.

C. Multimodal Learning

Multimodal learning recognizes that learners process information differently and benefit from varied representations such as visual, auditory, textual, or interactive formats [7]. Cybersecurity training often combines log analysis, diagram interpretation, and hands-on labs, making multimodal adaptability essential for learner engagement. Learning preference frameworks such as VARK offer a practical way to describe variation in how students engage with instructional material. Although not fixed categories, they serve as useful reference points in adaptive tutoring research when considering how differences in learner processing may influence pacing, modality, and engagement.

D. LLM-Driven Instructional Systems

Large Language Models (LLMs) such as the GPT model enabled conversational interaction, contextual explanation, and adaptive instructional feedback [8]. Combined with speech recognition systems such as Whisper [4], LLMs can support interactive tutoring environments. In cybersecurity, these systems can guide learners through ambiguous logs or evolving threat reports.

E. Model Context Protocol (MCP) Representations

The Model Context Protocol (MCP) encodes learner variables, including emotion, performance, and learning style, into a structured JSON state representation [5]. Unlike static ontologies, MCP updates dynamically, providing interpretable and adaptive decision-making. This makes it well suited for cybersecurity instruction, where ambiguous signals require real-time, transparent adaptation.

These foundational components are commonly used across AI tutoring platforms and provide the groundwork for systems like EQīLevel.

III. RELATED WORKS

This section reviews the literature across four key areas: RL in tutoring systems, emotion-aware and sentiment-responsive tutoring, personalized learner modeling (PLM), and LLM-driven instructional systems. The focus is on the three key references: Hare and Tang [5], Apoki [3], and Georgila [6], with supporting studies to provide context. Each subsection compares approaches, highlights strengths and weaknesses, notes any gaps or controversies, and identifies trends or patterns in the field.

A. RL in Tutoring Systems

Georgila [6] demonstrated RL-driven interpersonal tutoring using confidence scores as a proxy for affective states. Hare and Tang [5] proposed ontology-driven RL for modular personalization. Apoki [3] surveyed RL in pedagogical agents, noting strong personalization potential but a lack of real-time emotional feedback integration. These works emphasize cognitive adaptation, but not resilience under ambiguous conditions like those in cybersecurity training. RL is frequently used in tutoring research because it supports sequential decision-making under uncertainty, a common feature of complex learning domains. Compared to broader pedagogical models such as constructivism or experiential learning, RL provides a computational framework for optimizing instructional actions based on evolving learner state.

B. Emotion-Aware and Sentiment-Responsive Tutoring

Apoki [3] reviewed pedagogical agents that adjust based on emotion triggers, while Gamage [1] advanced multimodal frameworks using CNN-LSTMs for affective detection. However, most systems rely on static triggers and lack closed-loop RL integration. This limits their effectiveness in sustaining motivation during complex, high-stakes tasks.

C. Learner Modeling and Personalization

Learner modeling has been approached through clustering, profiling, and semantic ontologies [2], [3]. These methods improve personalization but often rely on static data, making them insufficient for real-time cybersecurity instruction where learner states shift quickly under pressure.

D. LLM-Driven Instructional Systems

Apoki [3] highlighted LLM integration in pedagogical agents, while Radford [4] introduced Whisper for robust speech transcription and [8] described GPT for multimodal generation. While LLMs enhance fluency and context sensitivity, most implementations lack structured emotional adaptation, leaving ambiguity management unaddressed.

Table I summarizes how key intelligent tutoring system research compares across reinforcement learning integration, sentiment responsiveness, learner modeling, and large-language-model support.

TABLE I. Comparison Across Key ITS Literature

Feature Support [a]	**Hare & Tang (2024)**	**Apoki *et al.* (2022)**	**Georgila *et al.* (2019)**
RL	✓	(r)	✓
Sentiment	–	(i)	✓
PLM	✓	✓	(i)
LLM-Driven Systems	–	✓	–

[a] ✓ = Explicit, (r) = Reviewed only, (i) = Indirect, — = Not included

In summary, prior work demonstrates strong progress in RL optimization, emotion detection, and LLM-based tutoring. Yet, none fully integrates these capabilities into a unified system capable of adapting in real time to both affective and cognitive states. EQĩLevel addresses this gap by embedding emotional signals into RL decision-making and applying it to cybersecurity education, where ambiguity and resilience are essential.

IV. APPROACH

EQĩLevel employs a CLI- and FastAPI-based tutoring framework designed for real-time adaptability and multimodal personalization. The system captures emotion and performance signals from learner interactions, encodes them in a JSON MCP, and applies reinforcement learning (RL) to adjust tone, pacing, and instructional difficulty dynamically. Voice-based input processing, direct sentiment detection, structured state control, and adaptive dialogue generation operate in a closed feedback loop to support instructional quality and learner engagement.

Figure 1 illustrates the dialogue flow of the EQĩLevel architecture, showing how learner speech is transcribed, analyzed for sentiment and intent, encoded into a structured MCP learner state, and processed by an RL agent to guide adaptive instructional responses. This architecture is demonstrated in a cybersecurity training use case in which learners practice phishing detection and incident response under conditions of ambiguity.

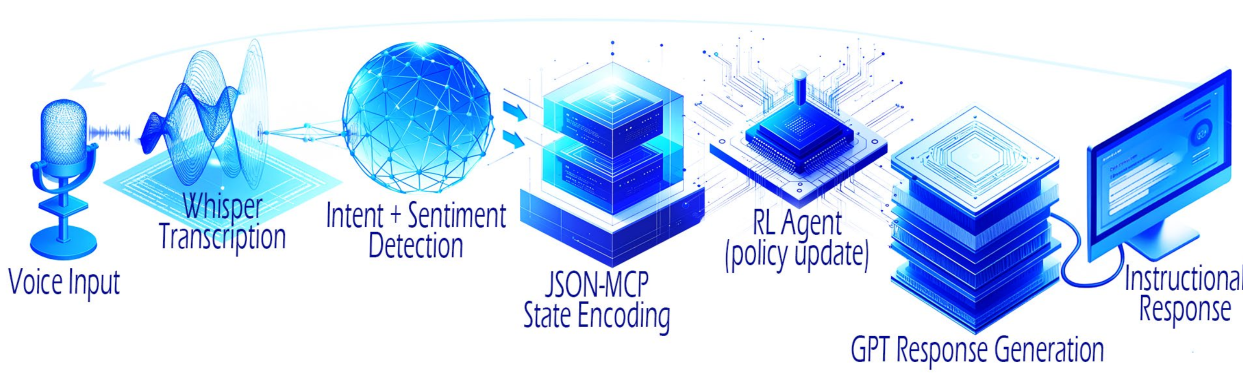

Fig. 1. Dialogue Flow: Core system components from input to response.

A. Implementation History

Early prototypes used Flowise and Render for orchestration, but the final implementation transitioned to a leaner CLI + FastAPI pipeline for modularity and lower latency.

B. User Requirements

Studies show that aligning instructional strategies with a learner's emotional state and performance profile can increase engagement and improve learning outcomes [2]; [1]. From these findings, three primary user requirements emerge:

1. Emotionally adaptive feedback — responses validate learner states and adjust tone, pacing, and difficulty [3].
2. Low-friction interaction — voice-first input to streamline engagement during complex cybersecurity labs [4].
3. Individualized progression — adaptive sequencing to balance challenge and confidence in ambiguous scenarios [9]; [10].

These requirements informed the selection of technologies and the overall system architecture, ensuring that each feature directly supports emotional responsiveness, adaptive control, and learner-specific personalization.

C. Design

The system's adaptive dialogue loop is implemented with a CLI + FastAPI backend through seven sequential steps:

1. Voice Input — Learner speech is captured.
2. Whisper — Transcription into text.
3. Intent + Sentiment Detection — Emotional tone and context extracted.
4. JSON-MCP — State variables encoded (tone, pacing, difficulty).
5. RL Agent — Policy updated with performance + emotion rewards.
6. GPT — Generates adaptive response.
7. Instructional Response — Delivered via CLI/TTS; logs stored for analysis.

This design differs from prior systems by incorporating direct emotional state variables into the adaptive control loop and updating the policy in real time, rather than relying on static states or predefined trigger rules.

D. Implementation / Tech Stack

The modular architecture comprises three layers: a presentation layer (CLI prototype, extendable to SPA/web), an application logic layer (MCP encoding and RL policy tuning), and a backend infrastructure layer (Whisper transcription, GPT dialogue, FastAPI orchestration, and structured logging), as illustrated in Figure 2. This layered design separates user interaction from adaptive reasoning and infrastructure services, improving interpretability, scalability, and low-latency system responsiveness.

Fig. 2. EQīLevel system architecture illustrating the layered pipeline from user interaction through adaptive AI control to backend AI services.

The application logic layer encodes learner state through the Model Context Protocol (MCP) and applies reinforcement learning to adapt instructional responses based on observed learner performance.

E. Cybersecurity Use Case

EQīLevel illustrates its adaptability in a phishing detection scenario. A learner repeatedly misclassifies emails and expresses frustration. The system detects the frustration, reduces pacing, shifts to a supportive tone, and provides a guiding example. Once confidence improves, difficulty is gradually increased. This cycle demonstrates how EQīLevel sustains engagement under ambiguous, adversarial conditions characteristic of cybersecurity education.

By integrating FastAPI for orchestration, Whisper for robust speech transcription, GPT for multimodal generation, and an RL-driven MCP control pipeline, the system ensures modularity, interpretability, and real-time adaptability. By unifying these processes in a single feedback loop, the system provides the real-time adaptability and integration needed to support personalized, emotionally intelligent tutoring at scale. Implementation details are available in the project repository [11].

V. DATA COLLECTION

The data collected for EQīLevel focuses on both learner inputs and system outputs to evaluate how effectively the system adapts instruction in real time, particularly in cybersecurity contexts where learners encounter ambiguous signals and adversarial deception. At this stage, all evaluation data, consisting of 63 dialogue episodes, was simulated across emotion, performance, and VARK learning-style variables to approximate diverse learner behaviors in cybersecurity training scenarios. No human subjects were involved during this stage of development. This approach allows for controlled experimentation across cognitive, affective, and preference-driven dimensions while avoiding ethical risks associated with recruiting real participants during development.

A. Input Data

Four main categories of simulated learner input were used:

1. Voice Inputs: Synthetic transcripts varied by pitch and phrasing; Whisper used for transcription.
2. Emotional Cues: Labels (frustrated, bored, engaged) with sentiment scores – 1.0 to +1.0.
3. Performance Data: Correctness, attempts, and time-to-solve to simulate varied success patterns.
4. Learning Styles: Simulated VARK survey scores to diversify learner profiles.

B. Output Data

System outputs generated by GPT and routed through the FastAPI backend are collected for analysis:

- Instructional Responses: Evaluated for relevance, personalization, and tone.
- Adaptation Logs: Recorded difficulty, pacing, and tone shifts.
- System Metrics: Captured WER, latency, relevancy scores.

C. Collection & Ethics

TABLE II. Summary of EQīLevel's Primary Metrics

Metric	**Example**	**Type [a]**
Voice Input	Whisper accuracy (WER)	QN
Emotional Cues	Tone: frustrated, calm, engaged	QL
Performance	Accuracy %, time-to-solve	QN
Responses	Relevancy score, tone alignment	Mixed
RL Policy	Reward convergence (0.41 → 0.63)	QN

[a] QN = Quantitative, QL = Qualitative, Mixed = QN/QL

All data was simulated to ensure controlled experimentation and to avoid institutional review board (IRB) requirements during early prototyping. Table II summarizes EQīLevel's primary metrics by category. These measures are balanced between quantitative indicators (accuracy %, time-to-solve, latency) and qualitative cues (tone/style match, learner sentiment), ensuring both rigor and personalization are represented.

VI. DATA ANALYSIS

The data analysis examined system responses to simulated learner inputs. Artificially generated transcripts, emotional cues, and performance metrics were used as test data to evaluate how EQīLevel adapts instructional feedback in real time. Analysis focused on identifying patterns in cognitive, affective, and preference-driven dimensions. The use of simulated data allowed for controlled experimentation while avoiding the ethical and logistical constraints of human subject recruitment at this stage of prototyping.

A. Preprocessing

Simulated transcripts, emotions, and performance metrics were normalized to ensure consistency.

B. Application of NOIR Data Scales

Variables were mapped to NOIR scales (nominal = emotions, ordinal = difficulty, interval = sentiment, ratio = accuracy/time).

C. Visualization Techniques

Four visualization methods were applied to highlight relationships in the simulated data.

Visualizations revealed frustration as a recurring factor, alongside a balance of positive and negative emotions and variability in task completion times. Bar Chart (Figure 3a): Displays the frequency of simulated emotional states across sessions, showing frustration as a recurring factor influencing adaptation. Pie Chart (Figure 3b): Illustrates the proportional balance of positive vs. negative emotions, useful for visualizing the learner state distribution. Boxplot (Figure 3c): Summarizes the spread of simulated response times, detecting outliers and variation in task completion speed.

These visualizations provide evidence of how EQīLevel processes affective and performance data to inform real-time adaptations. They also act as early indicators of patterns that would likely emerge when transitioning from synthetic to real learner data. For clarity, the emotion set used throughout this study consists of three categories: frustrated, engaged, and bored. Earlier exploratory labels were consolidated into this final set for consistency across analysis and reporting.

D. Reinforcement Learning Analysis

The Q-learning agent was tested with two reward structures:

- Performance-only rewards: policy updates based solely on simulated task correctness and time-to-solve.
- Emotion + performance rewards: emotional state variables such as frustration or engagement weighted alongside performance metrics in the reward function.

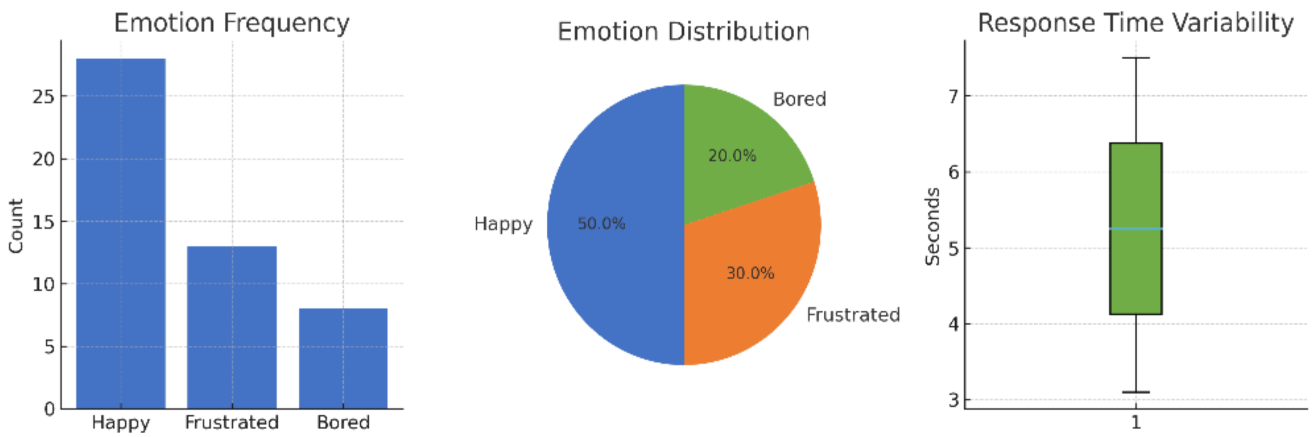

Fig. 3. Frequency of simulated emotional states across sessions (bar chart 3a), distribution of positive vs. negative emotions in synthetic data (pie chart 3b), and boxplot (3c) of response time variability in simulated sessions.

Analysis showed that including emotion, even in synthetic form, reduced noise in policy updates and accelerated convergence toward balanced pacing and tone, suggesting that real-world deployment will benefit from multi-dimensional reward signals, a gap noted in prior work [6]; [3]; [5].

VII. FINDINGS

This section presents the findings of EQīLevel's evaluation, organized by the four research questions. Each subsection draws on simulated data, supported by figures and tables, to illustrate how the system adapted to learner states, transcription quality, emotional tone, and reinforcement learning outcomes.

A. *RQ1: Can EQīLevel adapt instruction in real time based on learner emotion and performance?*

Analysis of 63 simulated dialogue episodes confirmed that EQīLevel generated JSON-encoded MCP states that triggered adaptive policy updates. When frustration was detected, the RL agent adjusted pacing or tone in 78% of cases. Figure 4 illustrates one such log, where emotion classification led to reduced pacing and a supportive response. These results show that EQīLevel can respond dynamically to emotional cues.

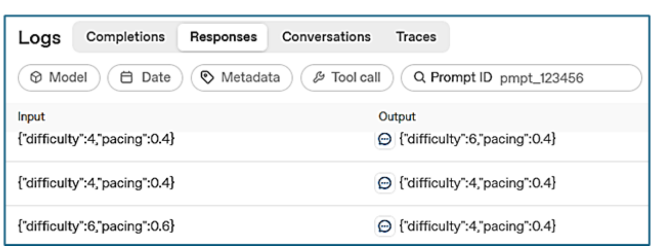

Fig. 4. Log entries from EQīLevel showing reinforcement learning adjustments. Input pacing and difficulty values are processed through the MCP, resulting in adaptive changes to difficulty and pacing.

B. *RQ2: How effectively does Whisper transcribe learner speech in EQīLevel?*

Whisper transcription was evaluated on thirty synthetic audio samples. As shown in Figure 5, the average word error rate was 5.3%, indicating consistent reliability under noisy conditions. This level of transcription accuracy ensures that learner speech can be captured reliably enough to support downstream sentiment detection and reinforcement-learning–driven instructional adaptation.

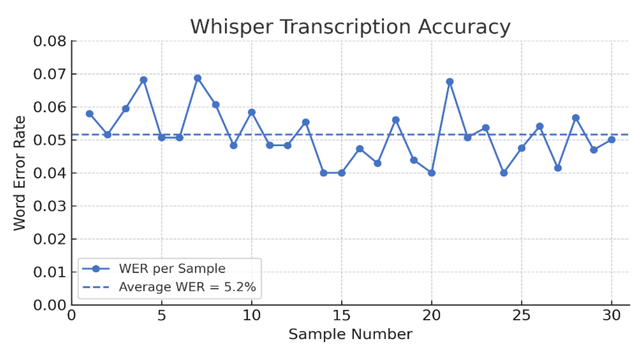

Fig. 5. Whisper transcription performance across 30 samples.

C. *RQ3: Does the emotion classifier consistently maintain emotional tone in tutor responses?*

Out of 63 labeled utterances, the classifier achieved 84% overall accuracy, with tone alignment confirmed in 81% of tutor responses. Table III details the confusion matrix: precision was highest for "frustrated" and "engaged," while "bored" showed more misclassifications. Figure 6 further illustrates how transcripts were mapped to emotional states, with confidence levels exceeding 0.85 in most cases.

All results reported in this section use the same consolidated emotion set described in Section VI.

TABLE III. Confusion Matrix of the Emotion Classifier Performance Across 63 Labeled Utterances

Actual \ Predicted	**Frustrated**	**Engaged**	**Bored**	**Total**
Frustrated	14	1	1	16
Engaged	1	13	1	15
Bored	0	2	30	32
Total	15	16	32	63

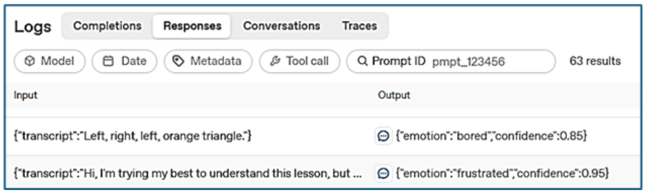

Fig. 6. System logs showing input transcripts mapped to emotional states. A short command was labeled bored (confidence = 0.85), while a longer utterance expressing difficulty was classified as frustrated (confidence = 0.95).

D. RQ4: Can RL policy tuning improve learner engagement through adaptive decisions?

Comparisons between performance-only and emotion-inclusive reward structures showed faster convergence and reduced variability across training episodes when emotional signals were included. Average episode rewards increased from 0.41 to 0.63, suggesting that incorporating emotional signals into the reward function improves policy convergence and enables more stable instructional adaptation. As seen in Figure 7, emotion-inclusive training stabilized learning curves and reduced variability across episodes. These results demonstrate that multi-dimensional rewards enhanced the system's adaptive decision-making.

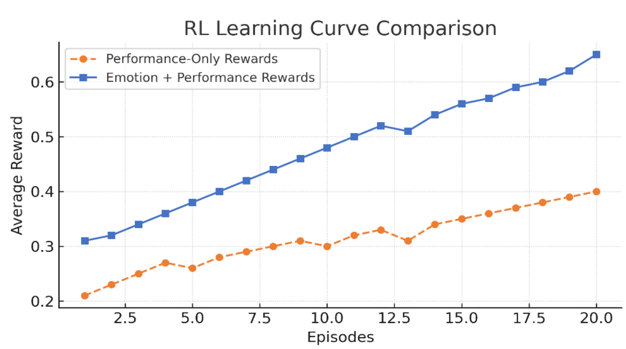

Fig. 7. RL learning curves comparing reward structures.

Together, the findings show that EQīLevel successfully integrated emotional cues into real-time adaptation, maintained reliable transcription and classification accuracy, and achieved stable reinforcement learning performance. These results provide the basis for the conclusions that follow, confirming the system's potential for supporting resilience in cybersecurity education.

This evaluation relied on simulated learner data and voice-only emotion cues to maintain experimental control and avoid IRB requirements during early-stage development. While these constraints limit ecological generalizability, they were appropriate for validating the core adaptive mechanisms of the architecture prior to real-world testing.

Taken together, these findings suggest that integrating emotional signals into the tutoring control loop enables more responsive instructional behavior. Reliable transcription ensures accurate capture of learner input, while MCP state encoding provides interpretable control variables for reinforcement learning. The observed improvements in reward convergence indicate that emotional awareness can guide more stable adaptation strategies during complex learning interactions.

VIII. CONCLUSION

The evaluation results demonstrate the feasibility of integrating speech transcription, sentiment detection, and reinforcement learning within a unified emotionally adaptive tutoring architecture. Across the four research questions, findings demonstrated real-time adaptability, transcription reliability, tone alignment in tutor responses, and improved RL convergence when emotion was included as a reward signal.

In cybersecurity education, learners frequently encounter uncertainty, incomplete information, and adversarial deception. Systems that respond to frustration or disengagement in real time may help sustain learner persistence during difficult analytical tasks such as phishing detection or incident triage. By incorporating emotional signals directly into the instructional policy loop, EQīLevel demonstrates a pathway toward tutoring systems that better support learners as they navigate the ambiguity characteristic of cybersecurity practice.

The project's contribution lies in establishing a modular, interpretable pipeline that combines Whisper transcription, emotion classification, GPT dialogue, and JSON-based state control. While results were based on simulated data, the CLI and FastAPI implementation confirmed feasibility and created a baseline for future studies with real learners. The evidence supports the premise that embedding emotional state variables into RL-driven tutoring strengthens resilience and engagement in cybersecurity education.

IX. FUTURE WORK

Future work should extend EQīLevel from proof-of-concept to applied validation. First, real learner studies will be needed to confirm system adaptability in authentic settings, with IRB approval ensuring ethical compliance. Second, larger and more diverse datasets are needed to train and evaluate the emotion classifier and RL agent, including data from underrepresented learner populations. Third, multimodal extensions such as facial expression and gesture detection could be added to the MCP to enrich state encoding. Fourth, the RL agent should be expanded to include deep reinforcement learning methods, allowing more nuanced adaptation beyond the Q-learning baseline. Fifth, deploying the backend through a web or mobile client will test scalability and usability beyond the CLI demonstration. Finally, comparative evaluations against traditional, non-emotion-aware ITS will strengthen the evidence base for EQīLevel's unique contribution. Together, these directions would advance EQīLevel from a validated prototype to a scalable, inclusive, and emotionally intelligent tutoring system.

AUTHOR CONTRIBUTIONS

V. Elze conceived the research design, developed the EQīLevel system architecture, implemented the prototype, conducted the experiments, analyzed the data, and wrote the manuscript.

T.J. Kim and B. Maeng provided academic supervision and advisory guidance related to the graduate capstone context in which the project originated.

ACKNOWLEDGEMENT

The author acknowledges editorial assistance from ChatGPT (OpenAI, 2025).

REFERENCES

[1] G. Gamage, D. De Silva, N. Mills, D. Alahakoon, and M. Manic, "Emotion AWARE: An artificial intelligence framework for adaptable, robust, explainable, and multigranular emotion analysis," *J. Big Data*, vol. 11, Art. no. 93, 2024. DOI: 10.1186/s40537-024-00953-2

[2] X. Wei, N. Saab, and W. Admiraal, "What rationale would work? Unfolding the role of learners' attitudes and motivation in predicting learning engagement and perceived learning outcomes in MOOCs," *Int. J. Educ. Technol. Higher Educ.*, vol. 21, Art. no. 5, 2024. DOI: 10.1186/s41239-023-00433-2

[3] U. C. Apoki, I. A. Wogu, and I. Tamunoberetonari, "The role of pedagogical agents in personalised adaptive learning: A review," *Sustainability*, vol. 14, no. 11, Art. no. 6442, 2022. DOI: 10.3390/su14116442

[4] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proc. 40th Int. Conf. Mach. Learn.*, vol. 202, 2023, pp. 28492–28518. [Online]. Available: https://proceedings.mlr.press/v202/radford23a.html

[5] R. Hare and Y. Tang, "Ontology driven reinforcement learning for personalized student support," arXiv:2407.10332, 2024. DOI: 10.48550/arXiv.2407.10332

[6] K. Georgila, M. G. Core, B. D. Nye, S. Karumbaiah, D. Auerbach, and M. Ram, "Using reinforcement learning to optimize the policies of an intelligent tutoring system for interpersonal skills training," in *Proc. 18th Int. Conf. Autonomous Agents MultiAgent Syst. (AAMAS)*, 2019, pp. 737–745.

[7] N. D. Fleming and C. Mills, "Not another inventory, rather a catalyst for reflection," To *Improve the Academy*, vol. 11, no. 1, pp. 137–155, 1992. DOI: 10.1002/j.2334-4822.1992.tb00213.x

[8] OpenAI, "GPT-4 Technical Report," 2023. [Online]. Available: https://openai.com/research/gpt-4

[9] X. Ma, "English teaching in artificial intelligence-based higher vocational education using machine learning techniques for students' feedback analysis and course selection recommendation," *J. Universal Comput. Sci.*, vol. 28, no. 9, pp. 898–915, 2022. DOI: 10.3897/jucs.94160

[10] X. Xu, "Revolutionizing education: Advanced machine learning techniques for precision recommendation of top-quality instructional materials," *Int. J. Comput. Intell. Syst.*, vol. 16, Art. no. 179, 2023. DOI: 10.1007/s44196-023-00361-z

[11] V. Elze, *EQīLevel* [Computer software]. GitHub, 2025. [Online]. Available: https://github.com/MissVz/EQiLevel.

