

CognisAI



DS4101: Human Technology Interaction

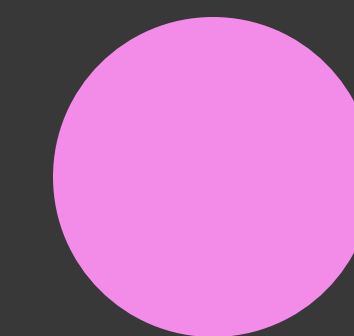
Ananya Shukla, Chaitanya Modi, Satvik Bajpai

Problem Statement



Students often experience unnoticed **cognitive overload, confusion, or disengagement** during self-directed digital learning. Current AI tutoring systems rely solely on **text-based interaction**, they cannot detect or respond to these fluctuating mental states, resulting in reduced learning effectiveness and increased frustration.

Why this Matters?



Effective learning depends on the learner's mental state, not just content delivery.

Current LLMs cannot sense when **learners are overwhelmed, disengaged, or confused**, leading to:

- Higher cognitive load
- Lower engagement
- Lack of personalized support

Literature Survey

State of Research: Gap Between Sensing and Reasoning

GazeTutor by D'Mello et al.	LLM Agents for Education by Chu et al.	NeuroChat by Baradari et al.
What It Is: A pioneering Intelligent Tutoring System (ITS) that was the first to use commercial eye-tracking for real-time disengagement detection.	What It Is: A framework defining modern LLM-powered pedagogical agents with advanced capabilities like memory, planning, and tool use.	What It Is: The first system to "close the loop" by integrating real-time biosensing (EEG) with a generative LLM to create a neuroadaptive tutor.
Key Features: Successfully monitored off-screen gaze (a proxy for "zoning out") and used simple dialogue prompts to reorient student focus, which showed a medium effect on improving deep reasoning.	Key Features: Moves beyond static, pre-scripted content to enable dynamic content generation. These agents can create personalized learning paths and perform automated error correction based on user's textual input.	Key Features: Proved that embedding a real-time physiological signal (EEG engagement score) into LLM prompts allows the tutor to dynamically adapt its tone and complexity, making the AI feel more "human-like".
The Gap: Sensed, but not Smart. It relied on a single, ambiguous behavioral signal (gaze) and couldn't distinguish boredom from confusion. Its content was static and its interventions were simple, pre-scripted prompts.	The Gap: Smart, but not Sensing. Despite their sophistication, these agents are blind to the learner's implicit state. They rely solely on textual interaction and cannot detect non-verbal, physiological cues like cognitive load, stress, or fatigue. They are powerful but reactive.	The Gap: A Narrow, Impractical Signal. While groundbreaking, it relies on a single, intrusive biosensor (EEG) impractical for scalable use. It is limited to one modality (text), missing the richer context from combining accessible signals like gaze and posture.

State of the Market: Personalized, but Not Present



What It Is: An AI-powered tutor using GPT-4 that provides Socratic questioning, hints, and personalized practice problems.

Key Features: Strong pedagogical framework with guided questions and adaptive practice based on performance.

The Gap: No Real-Time Awareness. It adapts to your answers, not your attention. The system is blind to cognitive load, stress, or waning engagement during a lesson.



What It Is: A premium tier featuring GPT-4 for conversational AI practice and an "Explain My Answer" feature.

Key Features: Simulates real-world conversation and provides adaptive difficulty paths based on user success.

The Gap: Blind to the Moment. It adapts based on past performance (post-hoc), not the user's real-time state. It cannot detect if a user is frustrated or disengaged right now.

Formative Study

Grounding CognisAI in User Reality: Formative Study



- Learners overwhelmingly **want personalized AI support** and already **rely on AI** during moments of **cognitive friction**.
- Many learners **abandon AI tools** when they fail to notice confusion or adapt explanations in real time.
- Learners want systems that can detect when they're stuck and **proactively guide** them with timely interventions.
- Real-time state-aware interventions significantly **reduce cognitive load** and **frustration** while improving performance.
- **Simplifying content** and **adapting tone** produce the strongest improvements in mental demand and frustration.
- Adaptive interventions enhance learning across **all modalities**, with each modality benefiting in different ways.

Solution

User State Classification



01

Baseline

Alert and ready.
Fresh start. Good
to go

02

Flow

Deeply engaged.
Everything clicks.
Learning feels
effortless

03

Struggling

Material is
challenging.
Working hard to
keep up. User
needs help

04

Overloaded

Mentally
exhausted. Can't
absorb more.
Need a break

Interventions



01

Baseline

- Normal content pacing
- Respond to user requests
- Monitor for state changes

02

Flow

- No modality switches
- No interruptions or suggestions
- Minimize UI changes
- Gradually increase complexity
- Only monitor for degradation

Interventions



03

Struggling

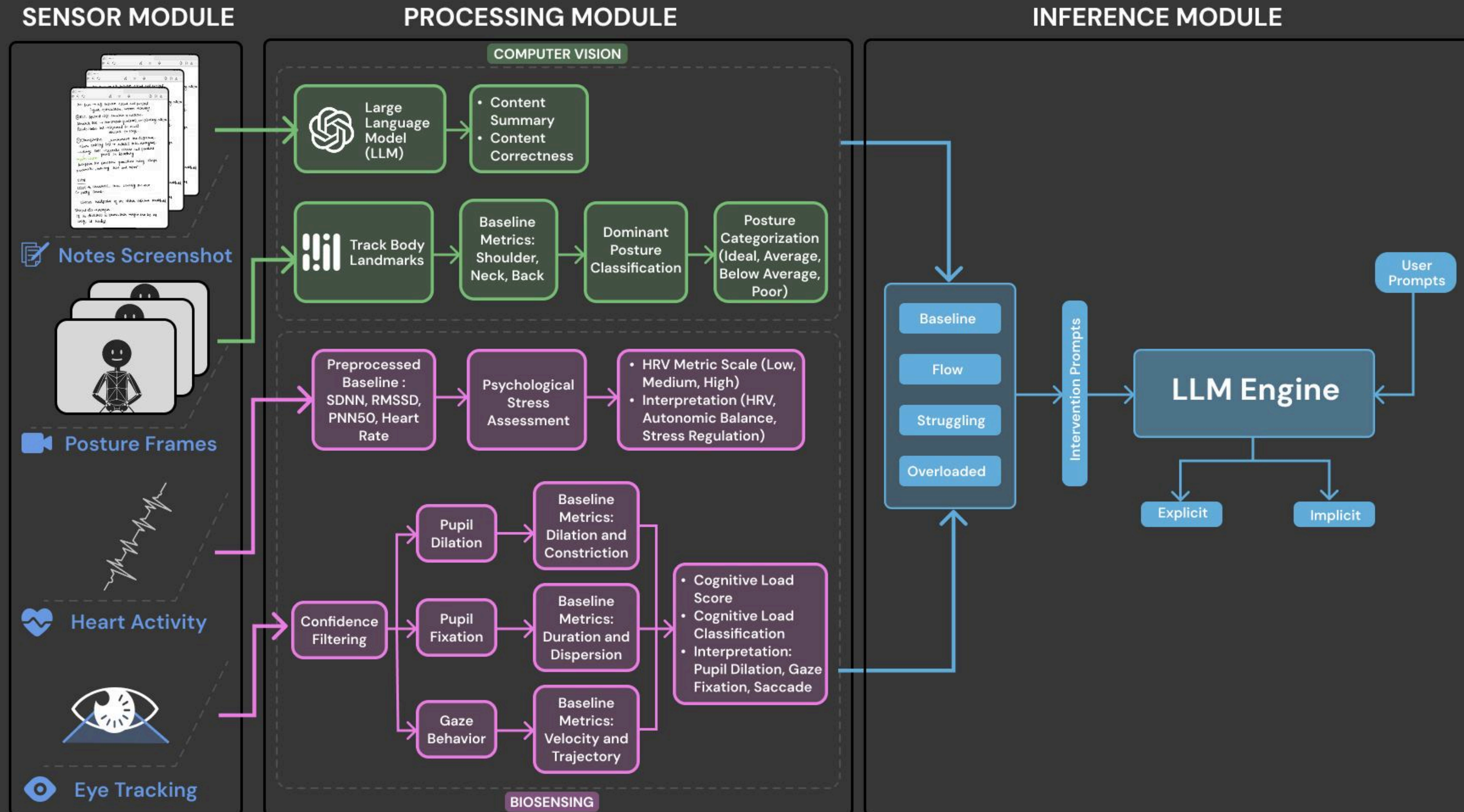
- Modality switch
- Content simplification
- Add examples: Visual diagrams, analogies, worked examples
- Pace adjustment: Slow down delivery
- Proactive suggestions

04

Overloaded

- Strong prompt to pause
- Breathing exercise: Guided relaxation
- Progress summary: Show what you've accomplished (encouragement)
- Pause content: Stop presenting new information

System Architecture



System Architecture

User Journey



Frontend

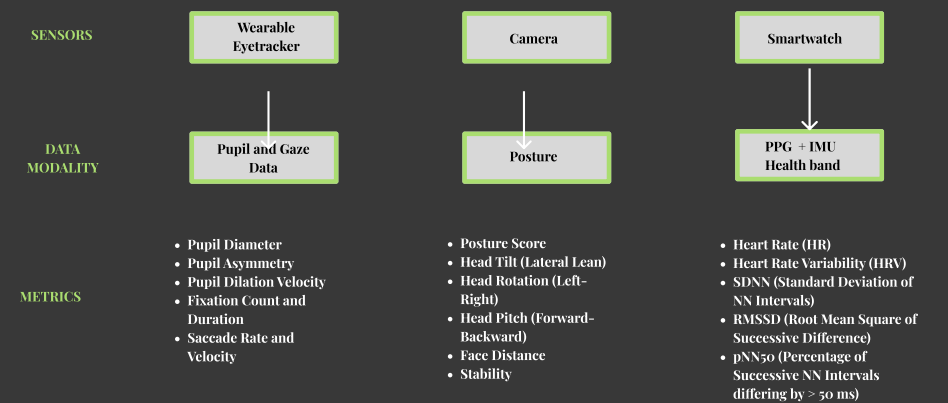


- Sensor Calibration
- Course Selection
- Learning Interface
- Settings

Sensor Connection

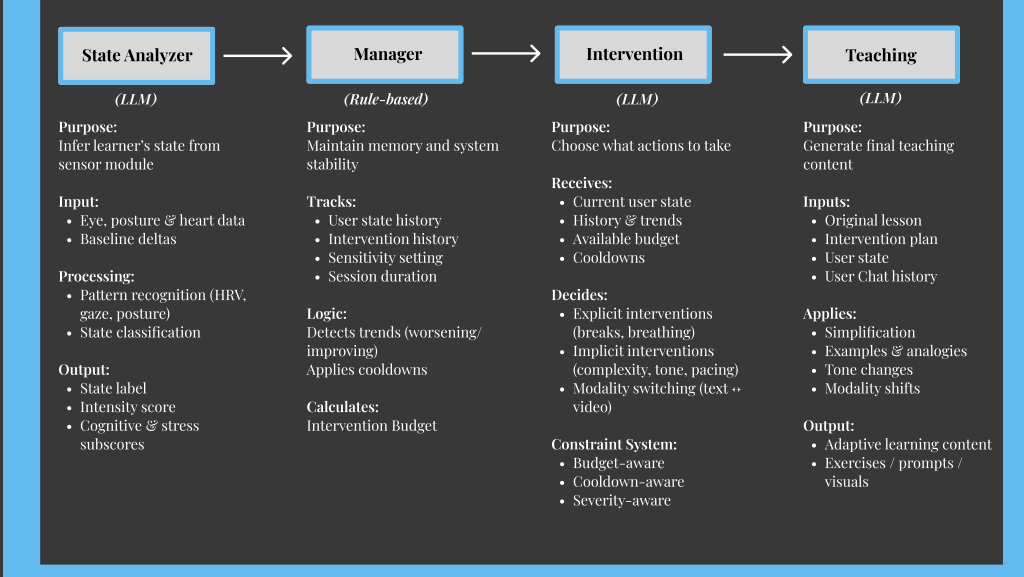
Course Catalogue, Interventions & LLM Responses
User Preferences & Context

Sensors



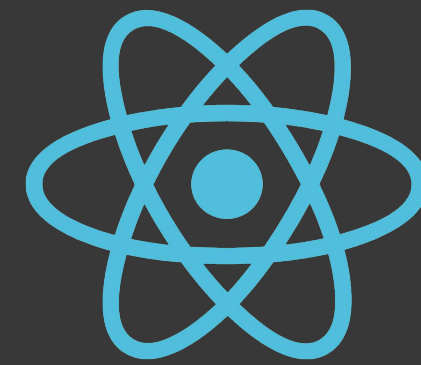
Processed Biosensory Data

Backend





Frontend

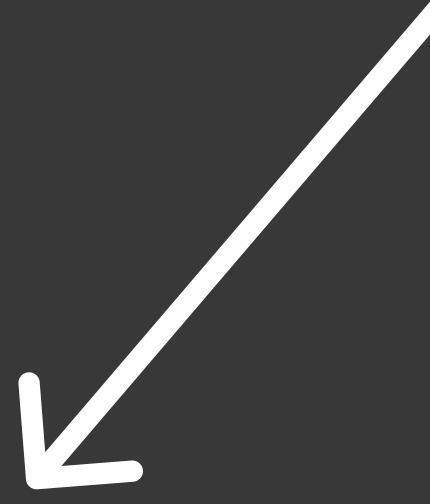


**Sensor
Calibration**

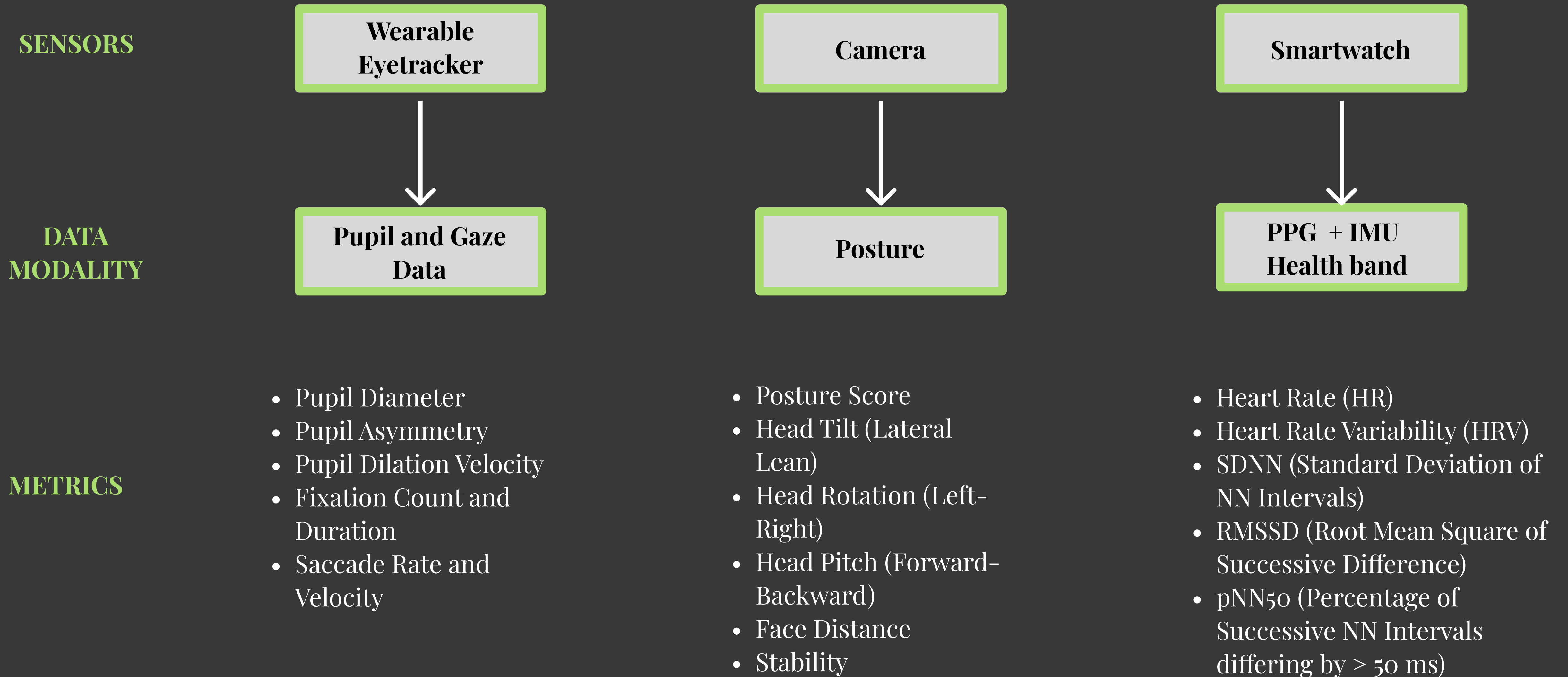
**Course
Selection**

**Learning
Interface**

Settings



Sensors



BASELINE (60s)

SENSORS

**Wearable
Eyetracker**

Camera

Smartwatch

**DATA
MODALITY**

**Pupil and Gaze
Data**

Posture

**PPG + IMU
Health band**

METRICS

- Pupil Diameter
- Pupil Asymmetry
- Pupil Dilation Velocity
- Fixation Count and Duration
- Saccade Rate and Velocity

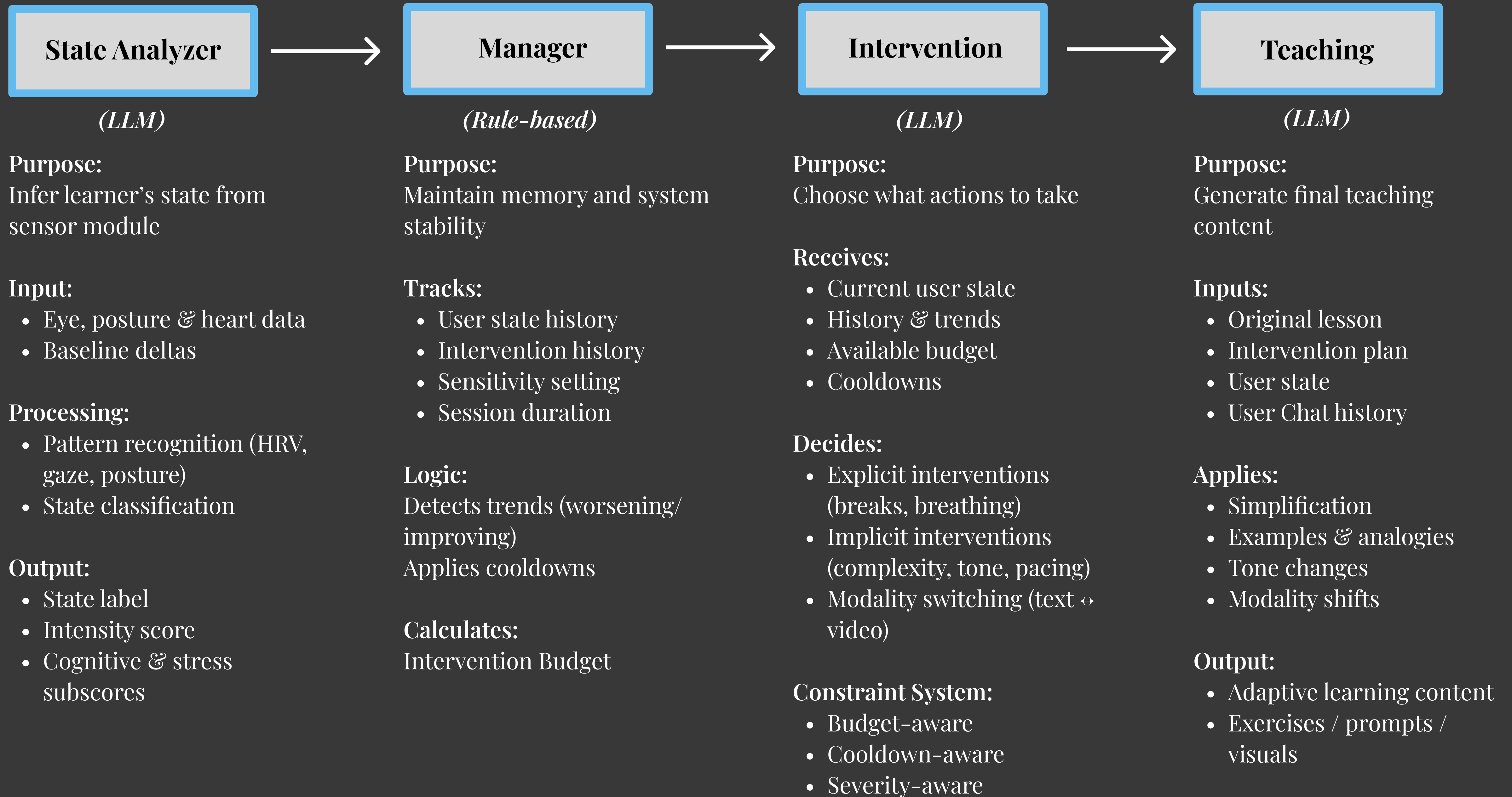
- Posture Score
- Head Tilt (Lateral Lean)
- Head Rotation (Left-Right)
- Head Pitch (Forward-Backward)
- Face Distance
- Stability

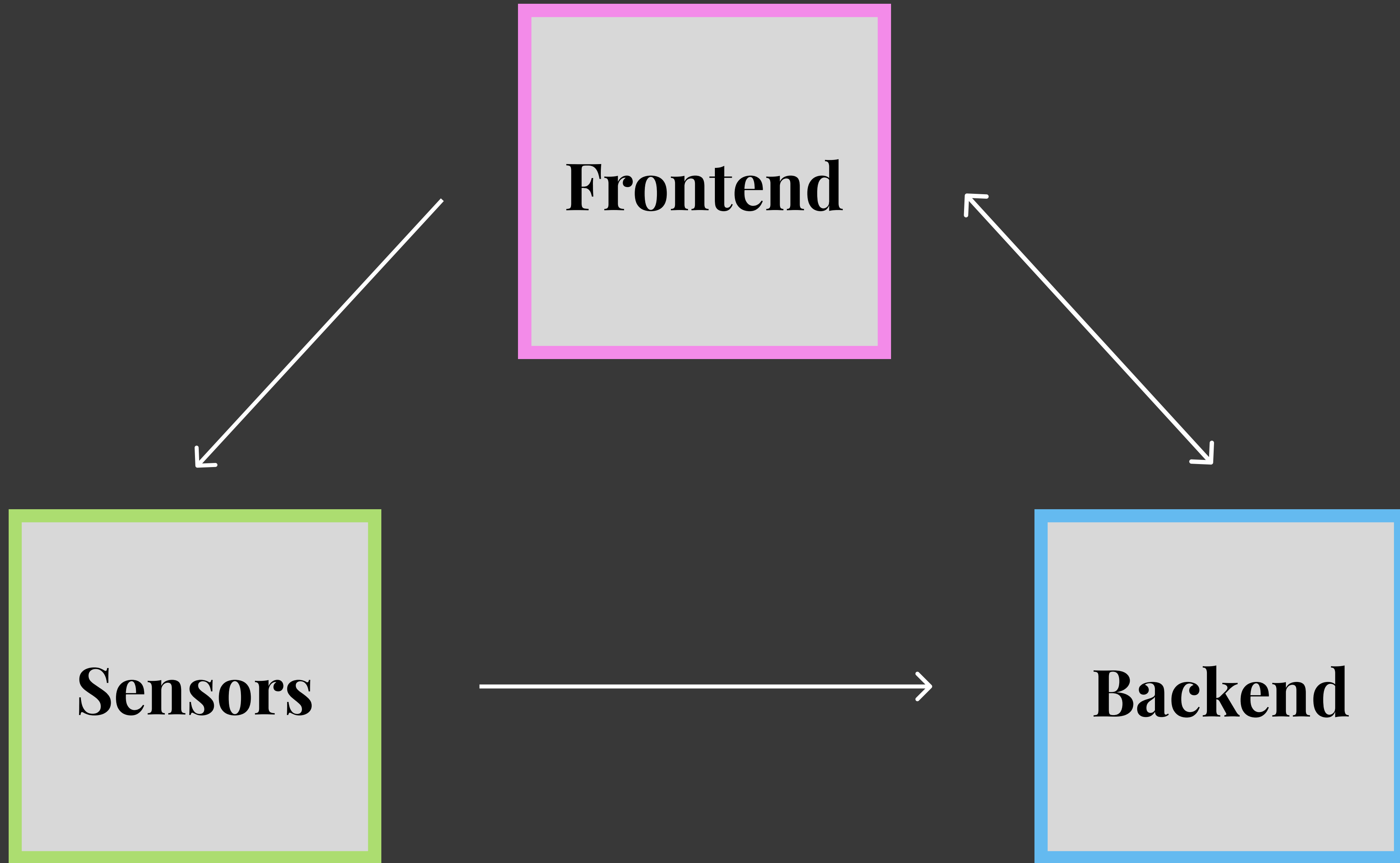
- Heart Rate (HR)
- Heart Rate Variability (HRV)
- SDNN (Standard Deviation of NN Intervals)
- RMSSD (Root Mean Square of Successive Difference)
- pNN50 (Percentage of Successive NN Intervals differing by > 50 ms)





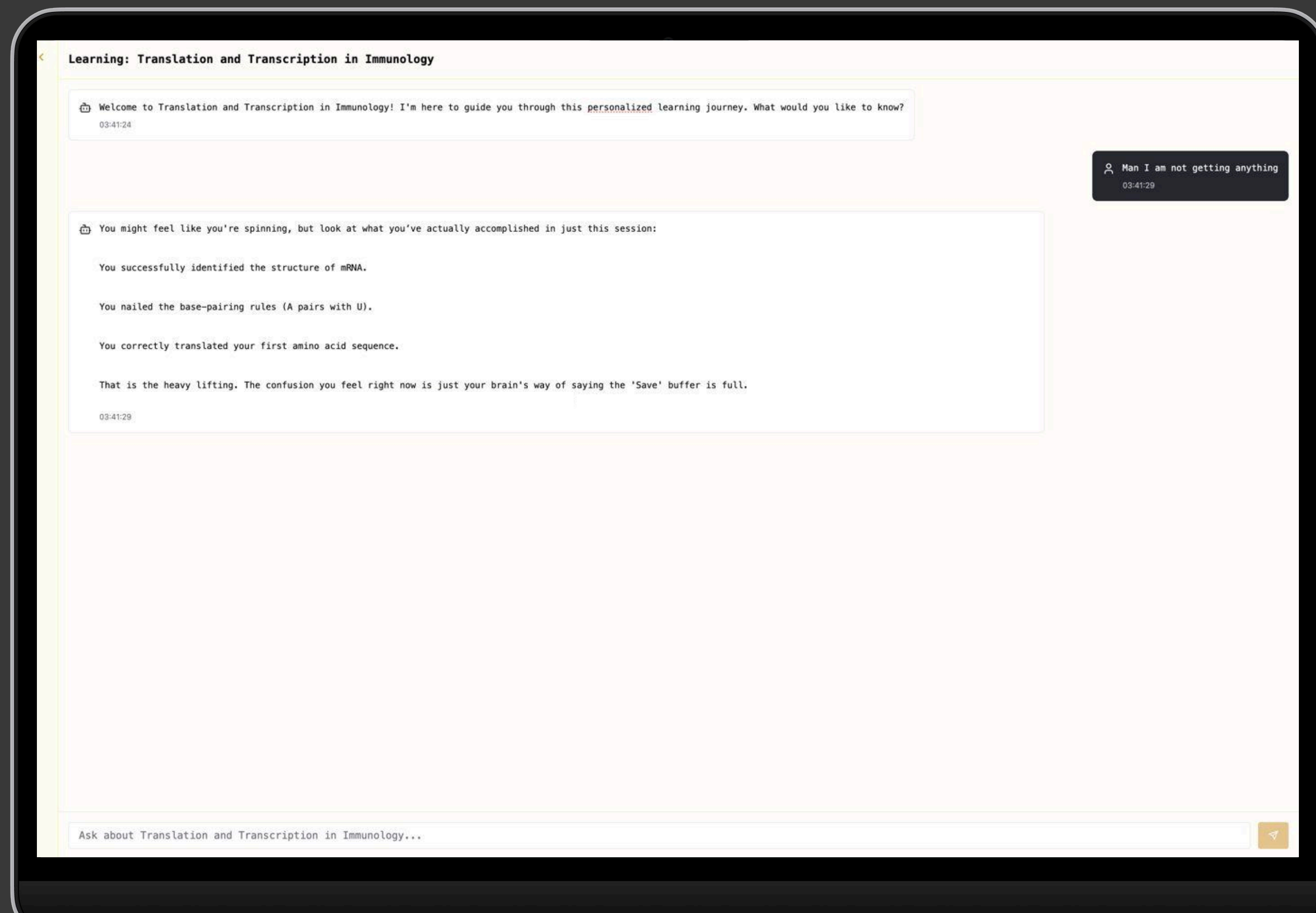
Backend





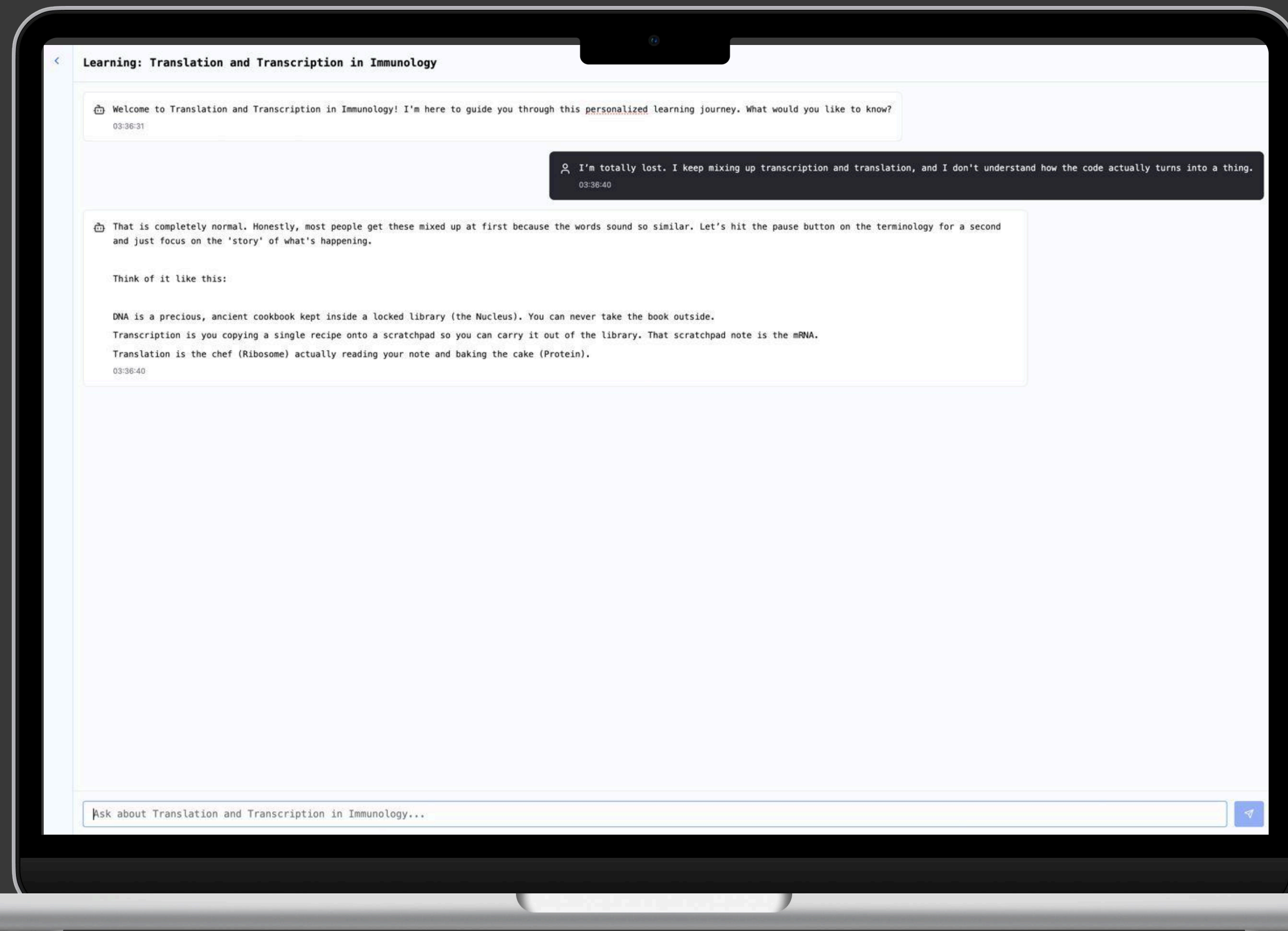
Learning Interface

Text Feedback + Supportive Messages: Adaptive guidance responding to your real-time learning needs



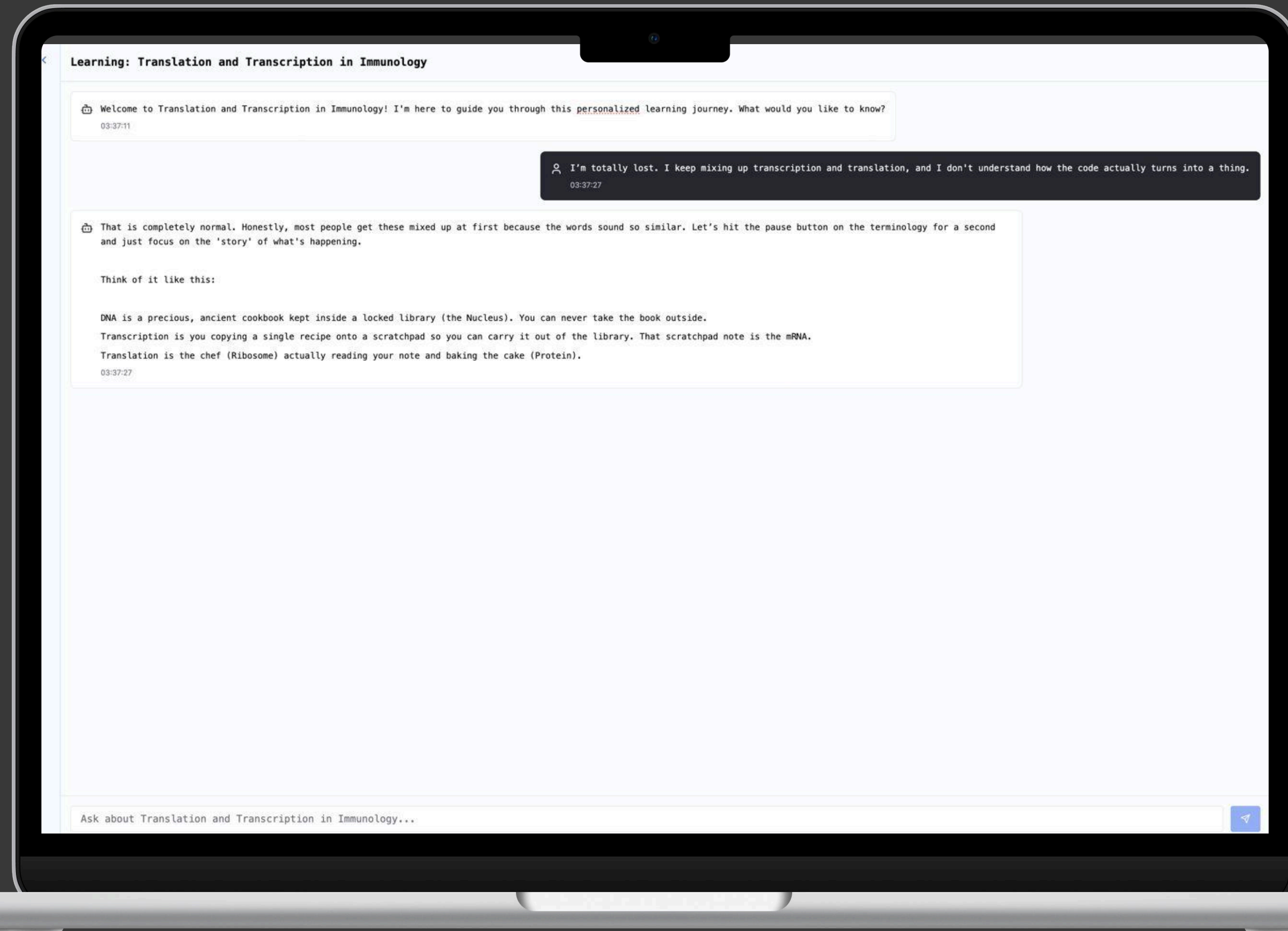
Learning Interface

Clarification & Explanation Support: Context-aware explanations that adjust to confusion or difficulty



Learning Interface

Simplified Explanations in Response to Struggle: Real-time simplification to enhance understanding



Modality Switch

Seamless modality shifts to improve comprehension through visuals

can you show me a video to teach this concept
14:57:10

Here's an excellent educational video about the immune system! This Kurzgesagt video explains "How The Immune System ACTUALLY Works" - perfect for understanding complex biological processes. The video is about 10 minutes long and uses amazing animations to make the concepts clear.
As requested
14:57:11

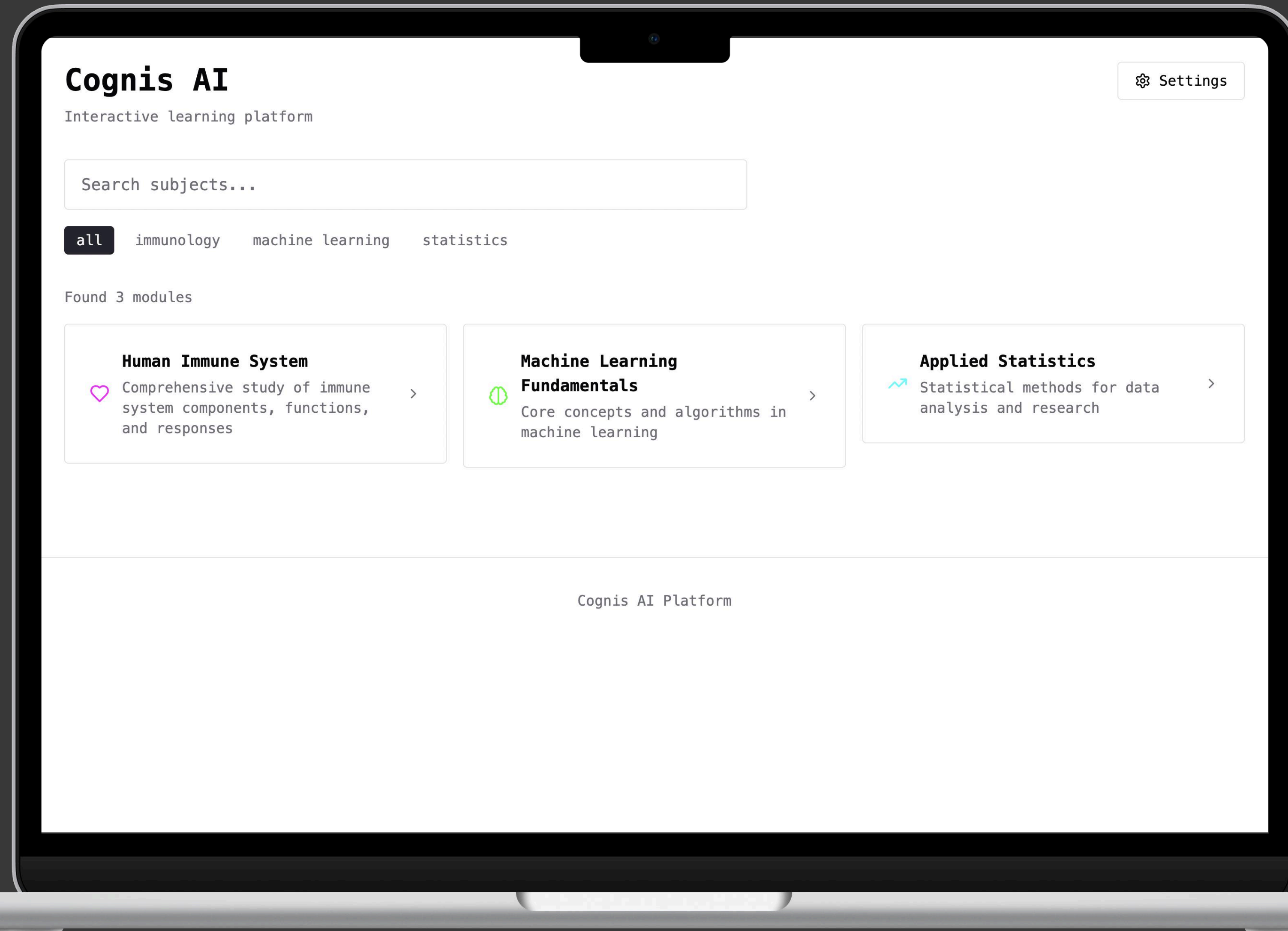
Video

How The Immune System ACTUALLY Works - IMMUNE
Kurzgesagt - In a Nutshell 10:48

▶ Play/Pause ⏪ Restart 🔍 Fullscreen

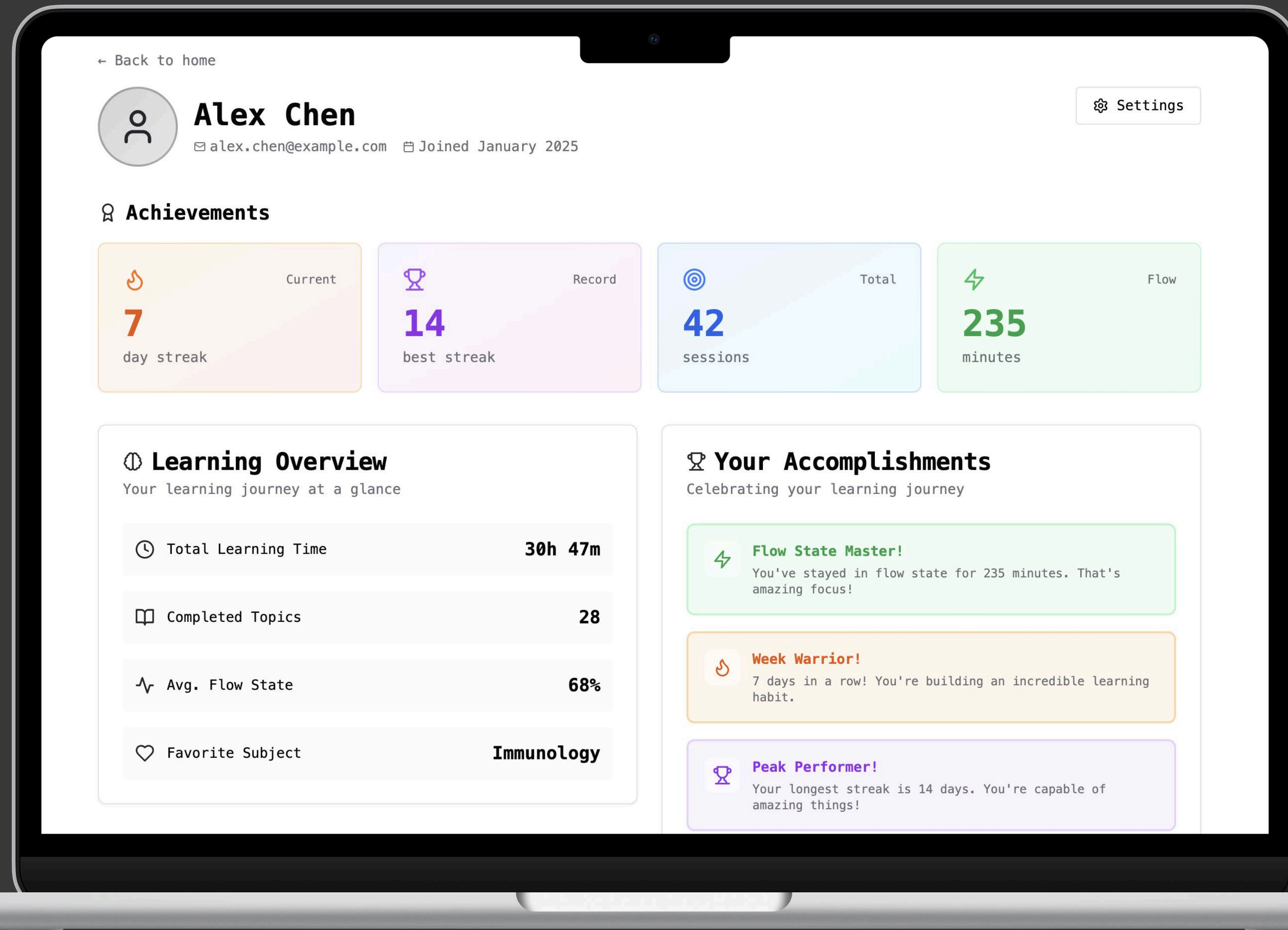
Course Catalogue

Discover subjects and choose your personalized learning module



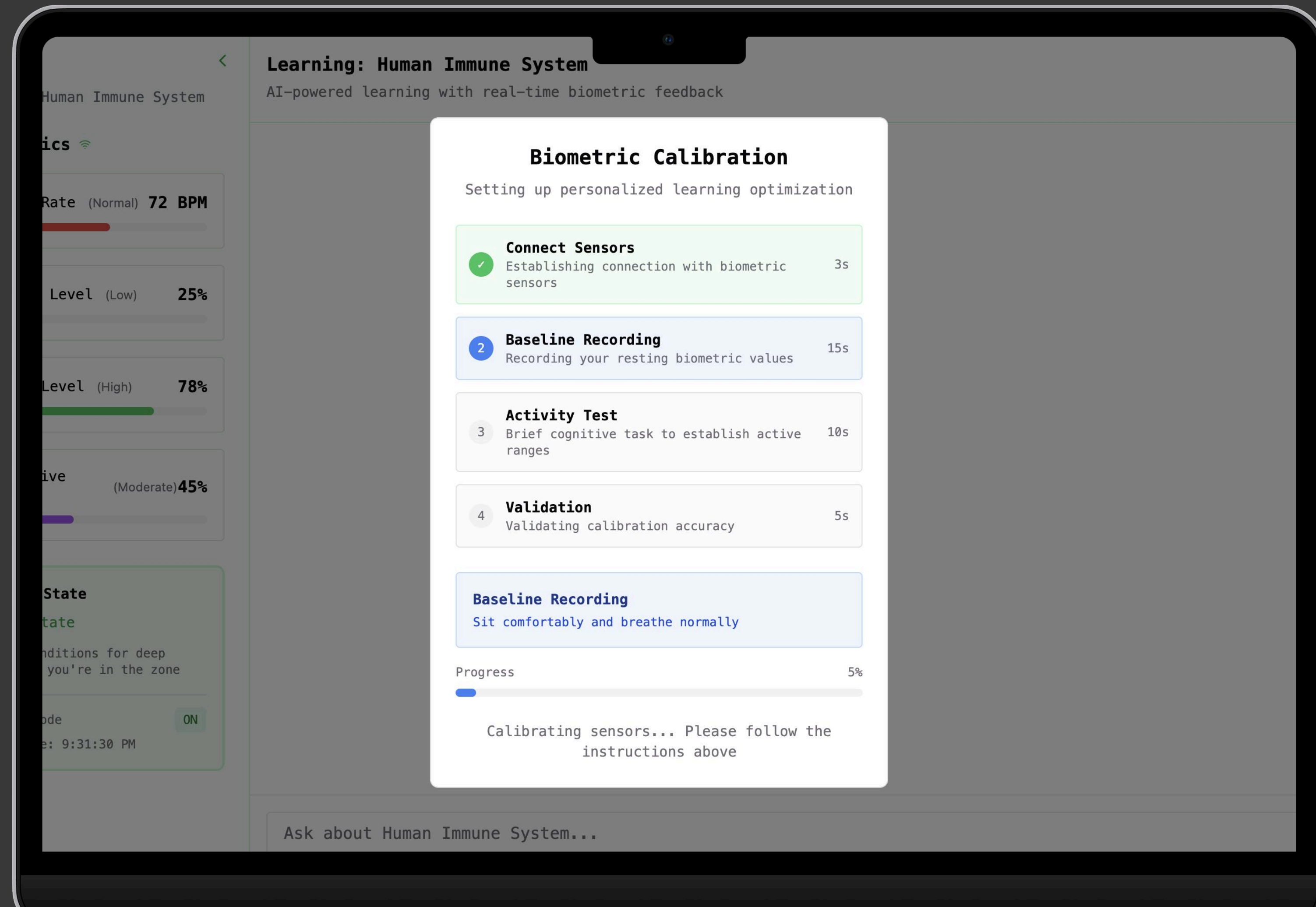
User Dashboard

Track your learning progress, achievements, and personalized insights



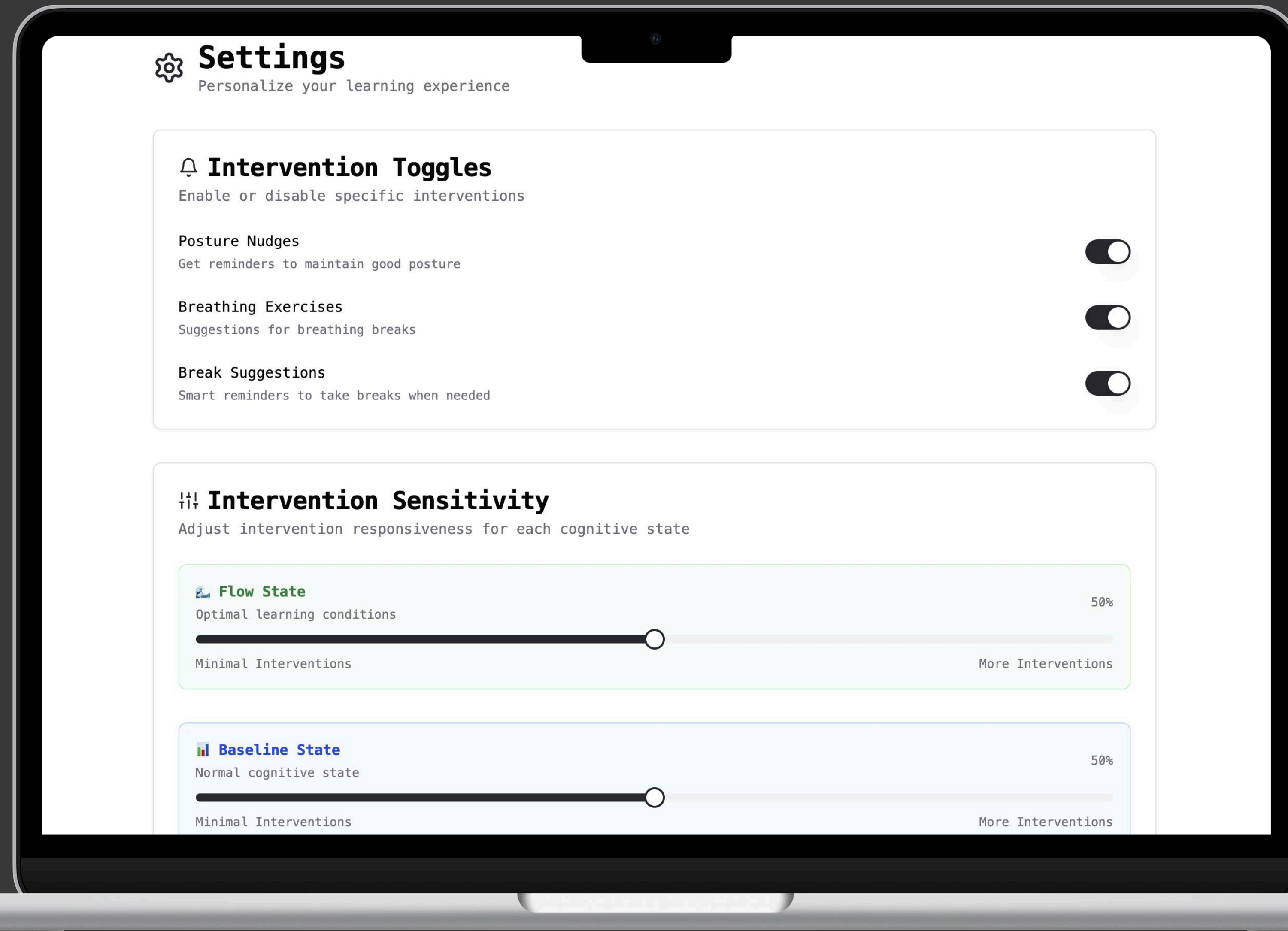
Biometric Calibration Flow

Setting up biometrics for optimized state-aware learning

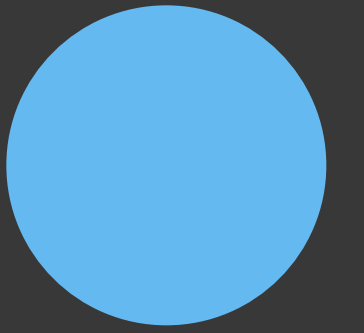


Settings: Intervention Controls

Customize intervention preferences and cognitive state sensitivity



Study Design



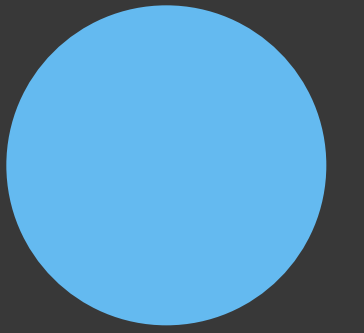
Objective

Evaluate the effectiveness of our real-time multimodal adaptation system, built from prior formative studies, in improving learning efficiency and user experience.

Participants & Task

- N = 11 Participants: Randomized order to minimize bias. (5 female, 6 male; M age = 21.2 years)
- Learning Topics: Users engaged in one of two complex modules:
 - Computer Science: Machine Learning (Support Vector Machines).
 - Biology: Transcription & Translation.
- Duration: Average session length of 15–20 minutes.

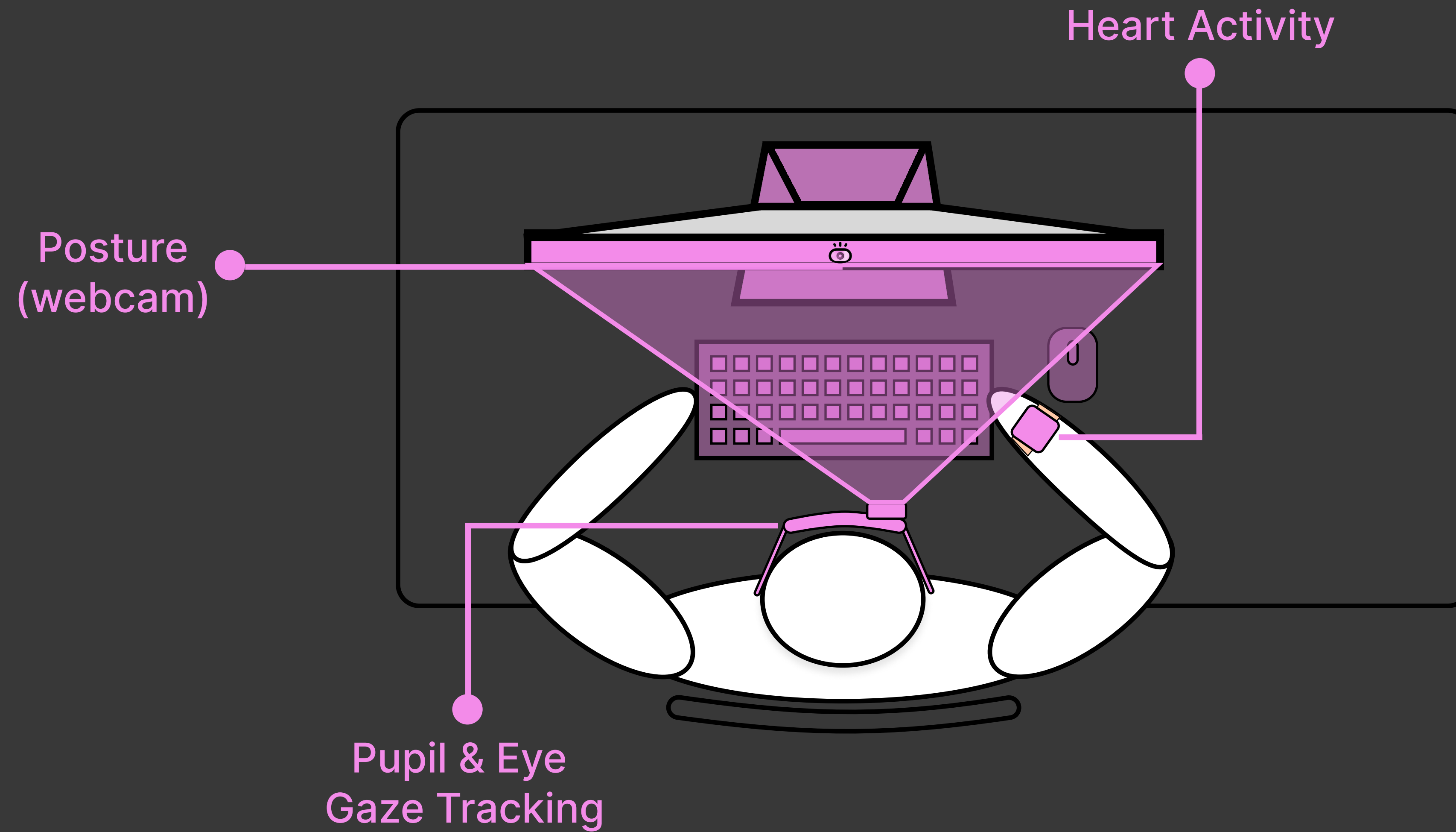
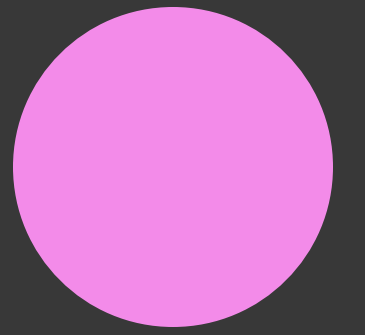
Study Design



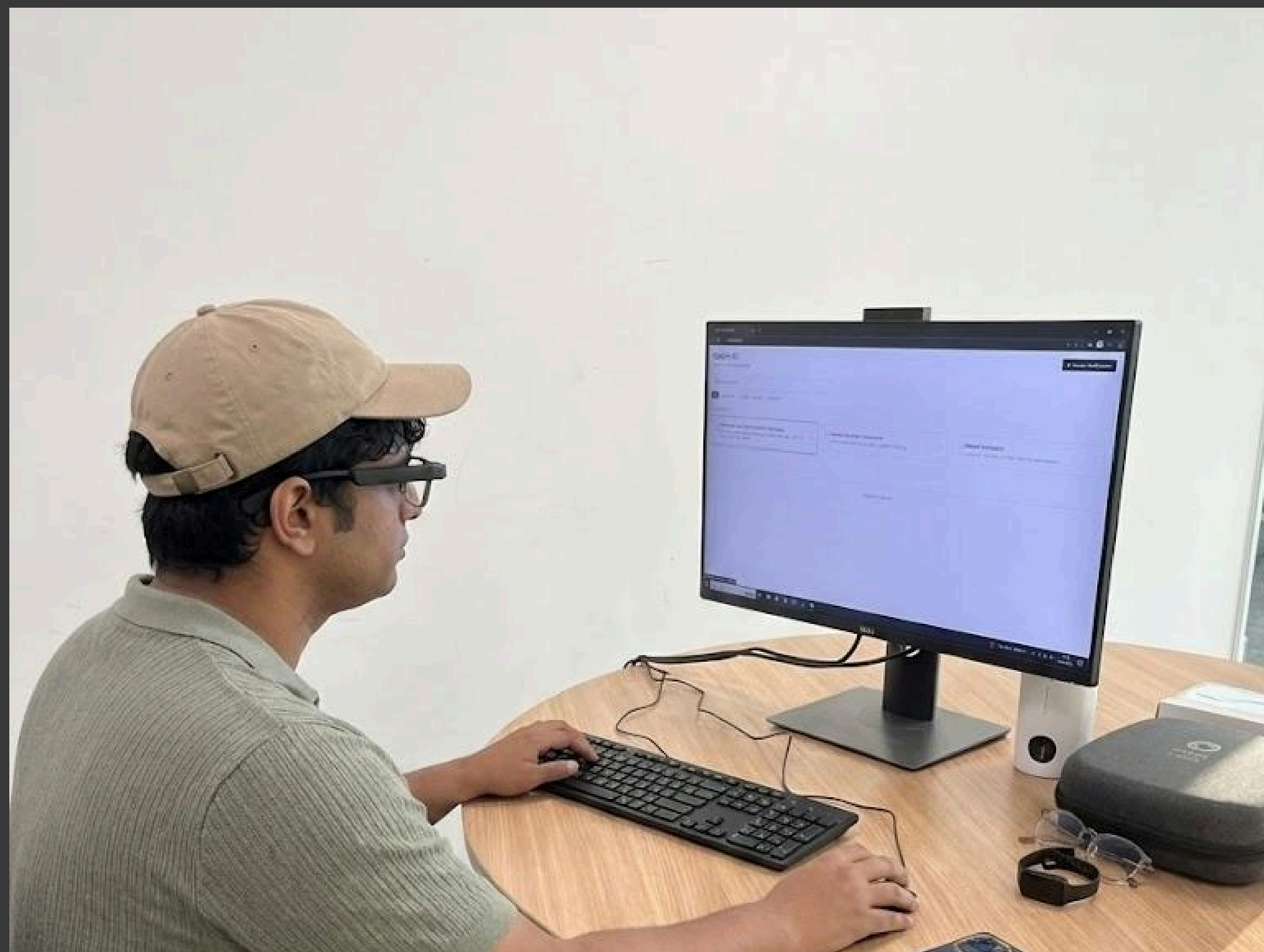
Experiment Protocol (The Flow)

1. Onboarding: Participants gave consent and were informed of the protocol; all physiological data was processed locally/anonymously.
2. Calibration: Baseline vitals recorded before starting.
3. Adaptive Task: Participants used CognisAI, which triggered interventions based on real-time states.
4. Assessment: A 5-question knowledge check followed by surveys (NASA-TLX and SUS).

The Setup



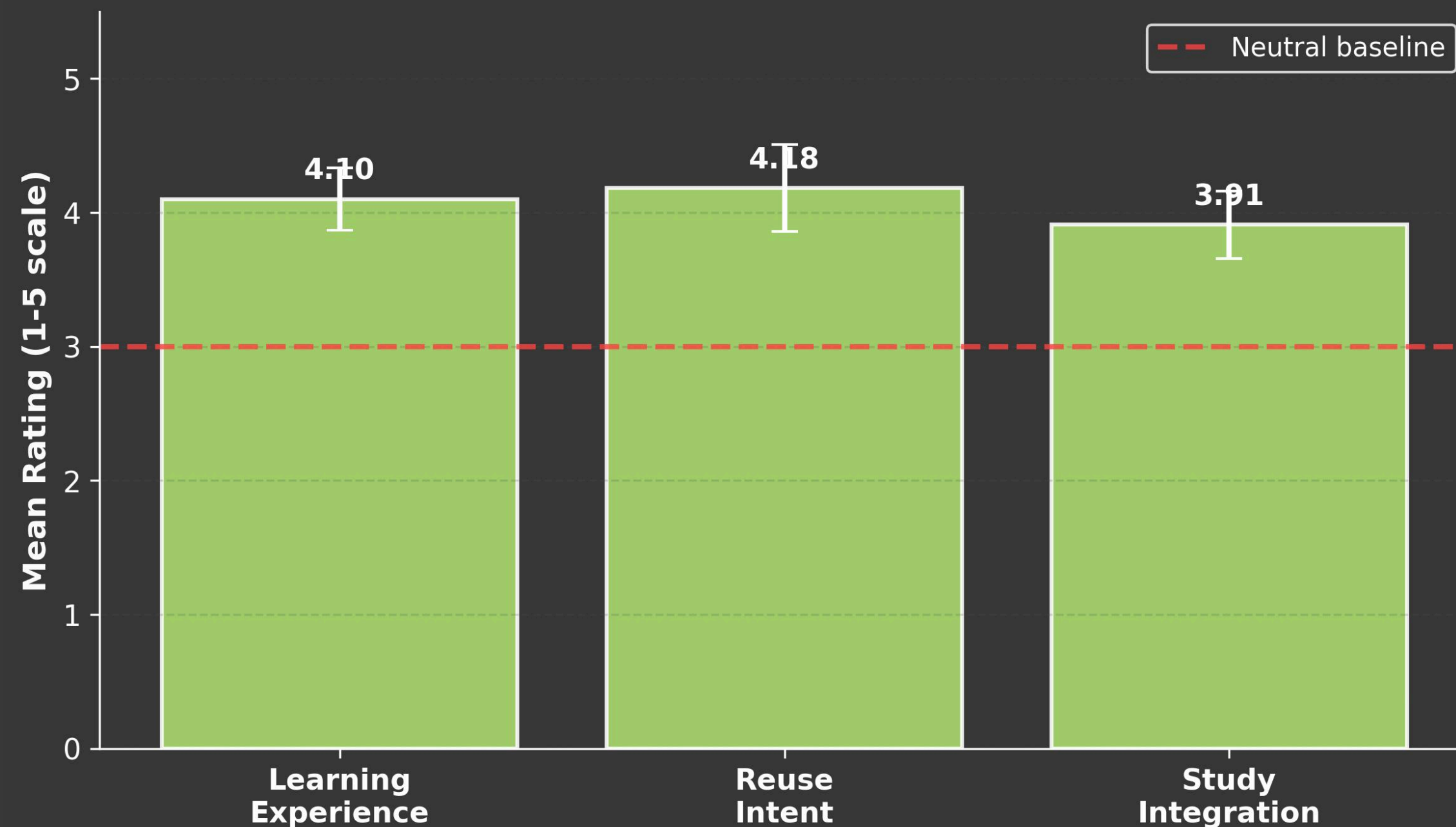
Experiment Pictures



Metrics & Validation

System Effectiveness

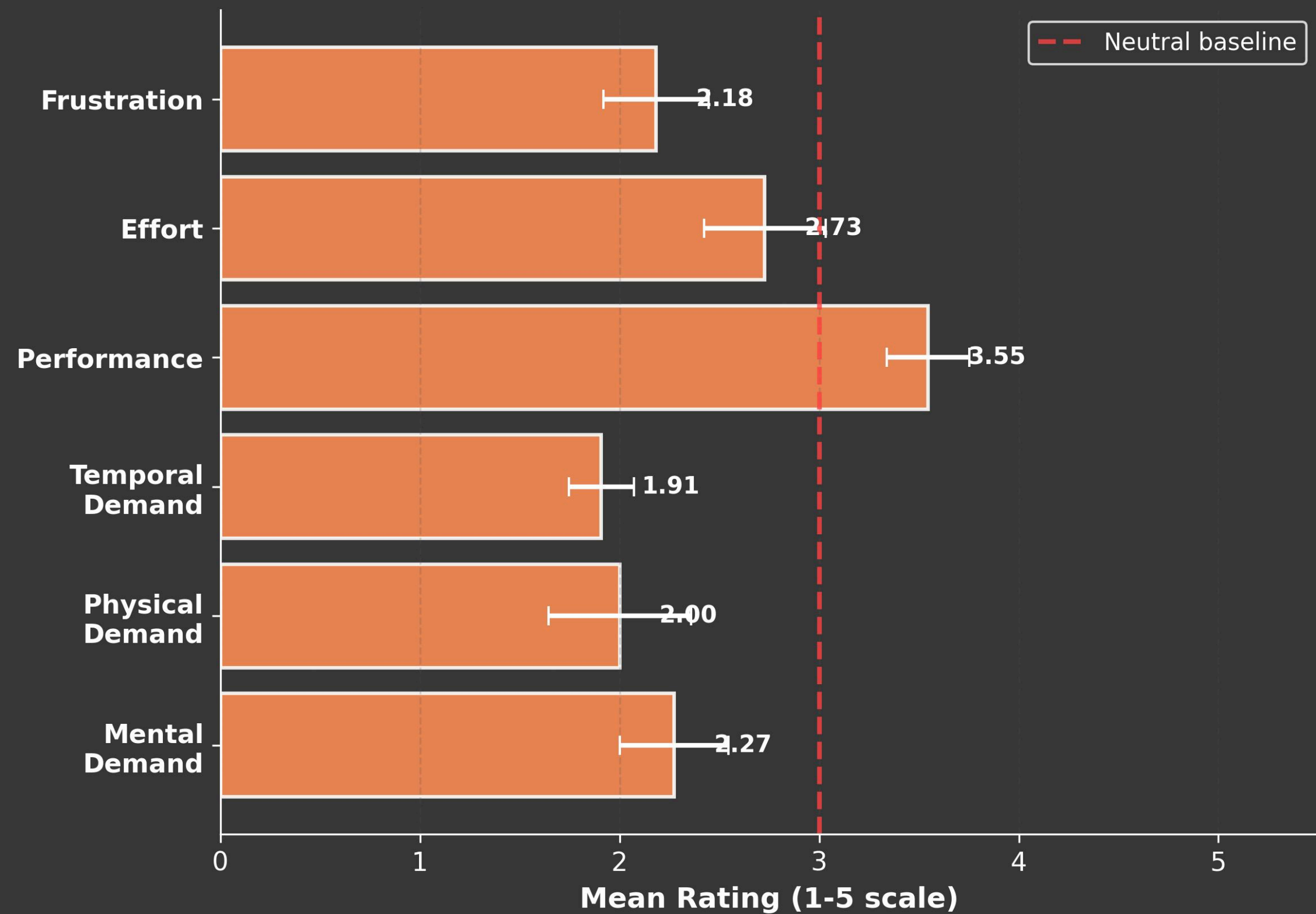
Figure 1. System Effectiveness: Key Learning Metrics (N=11)



- Users rated learning Experience rated 4.10/5, significantly above the neutral baseline.
- Reuse Intent was the highest metric (4.18/5), proving users want to use this again.
- Users rated Study Integration at 3.91/5, confirming the tool fits into daily routines.

NASA-TLX

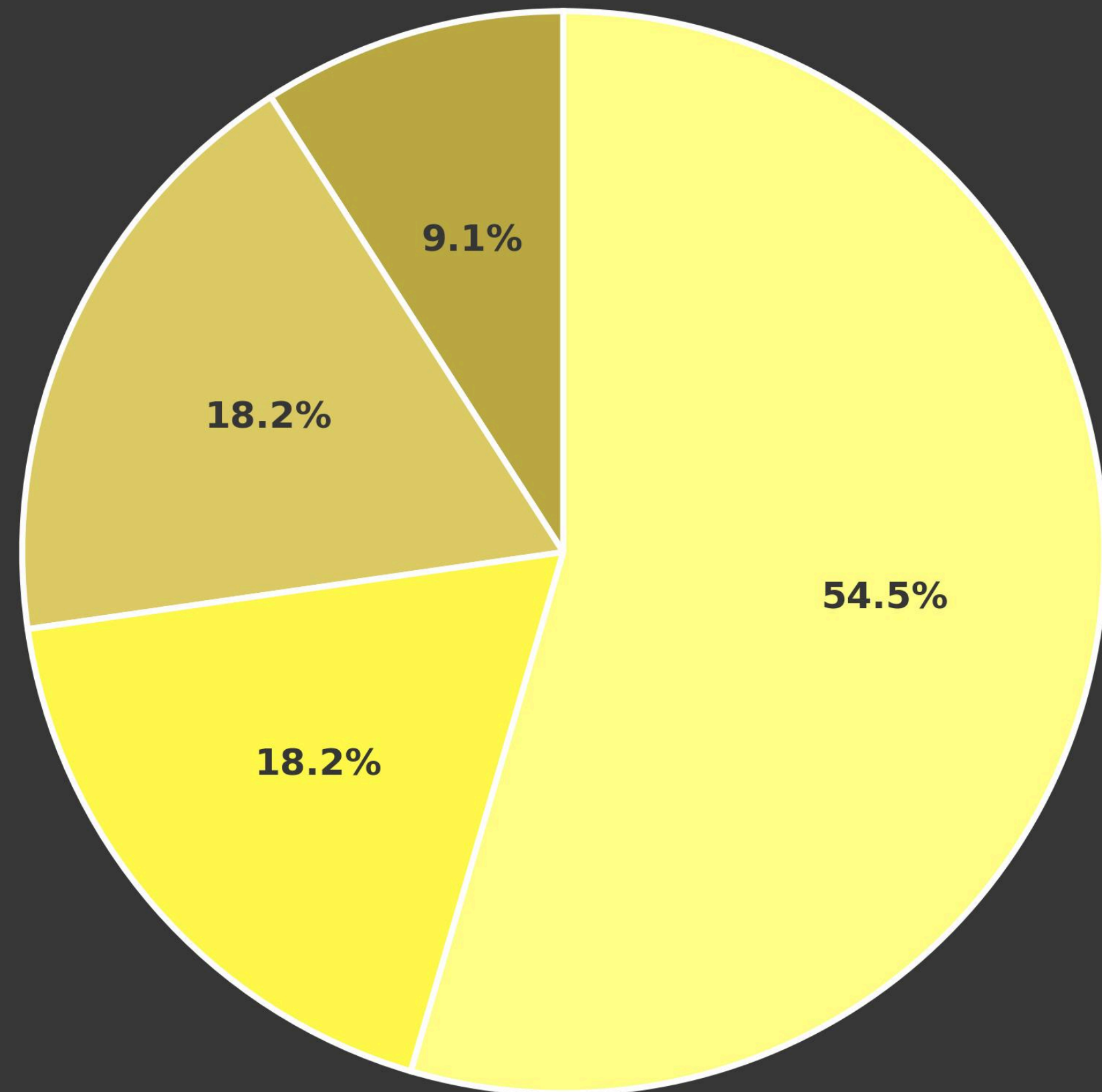
Figure 2. NASA-TLX Workload Assessment (N=11)



- **Low Frustration**, rated 2.18/5,
- **Users rated their own success at 3.55/5**, the highest metric on the chart.
- **Mental Demand (2.27) and Temporal Demand (1.91) were minimal, proving the system successfully managed cognitive load.**

Reuse Intent

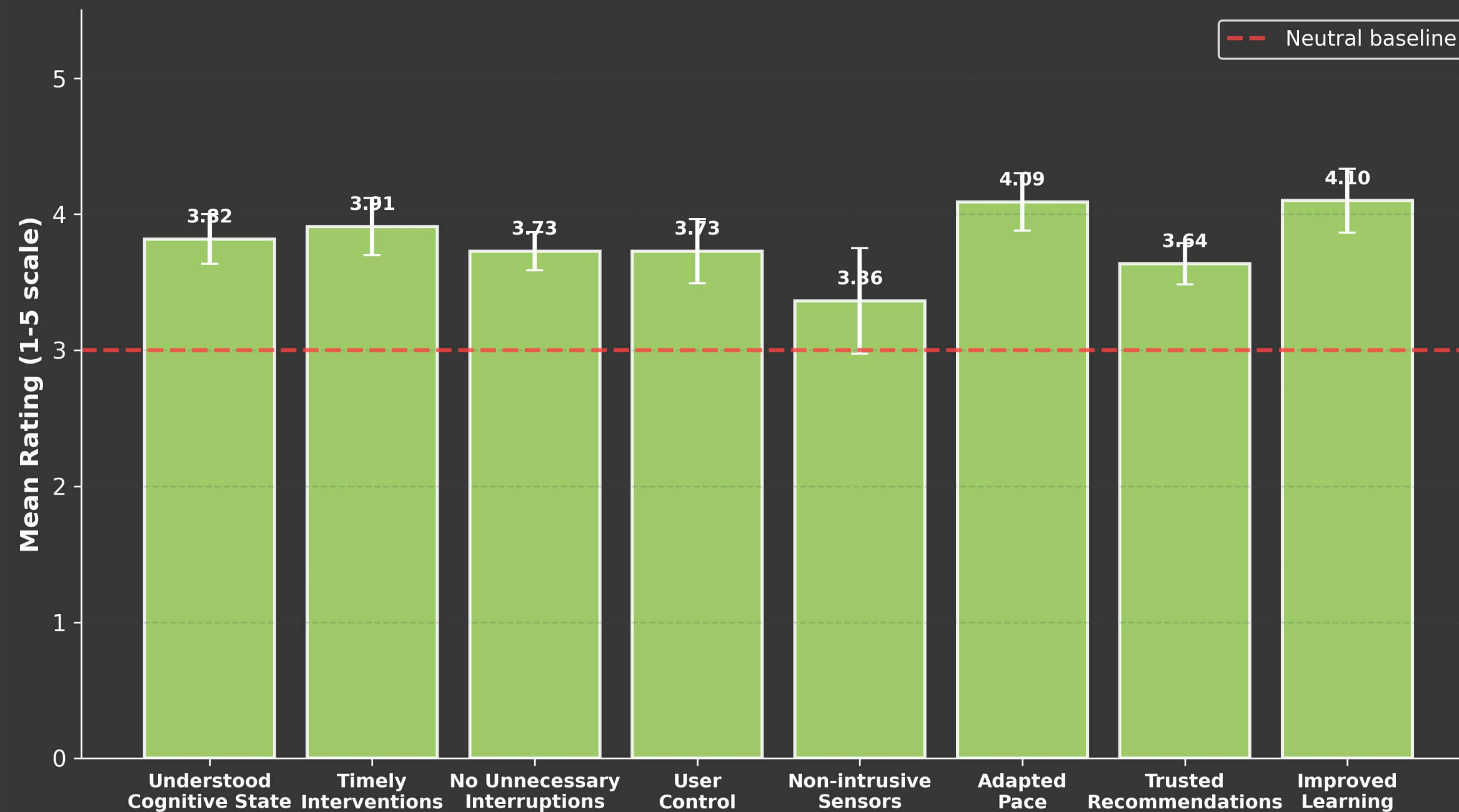
Figure 4. Distribution of Reuse Intent Ratings (N=11)



- **72.7% of users** rated their likelihood to reuse the system as "Likely" or "Very Likely."

System Usability Scale

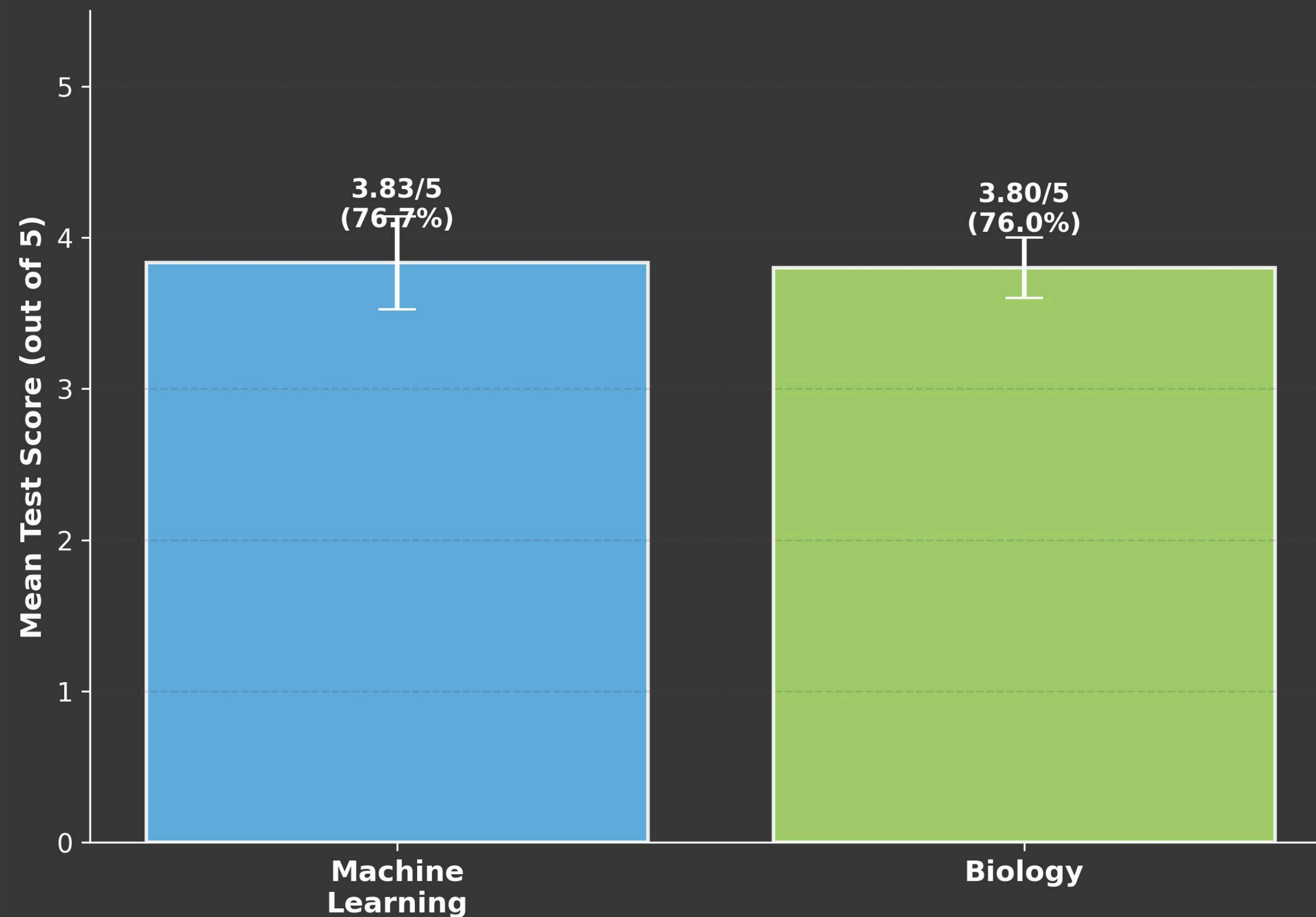
Figure 5. Trust and System Perception Dimensions (N=11)



- Users rated Improved Learning (4.10) and Adapted Pace (4.09)
- **Interventions were perceived as highly Timely (3.91), validating the state-detection algorithms.**
- **Non-intrusive Sensors (3.36) received the lowest score depicting friction with hardware setup.**

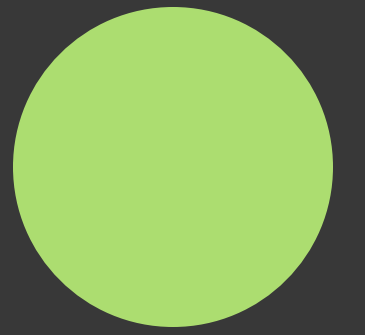
Objective Performance

Figure 2. Average Objective Test Scores by Topic



- **Performance averages (~76%) significantly exceeded the "Good" threshold of 60% established for the study.**

Challenges Faced



01

Intervention Timing

Ensure the system delivers the right help instantly when a user's state changes, eliminating any reaction lag.

02

Sensitivity Balance

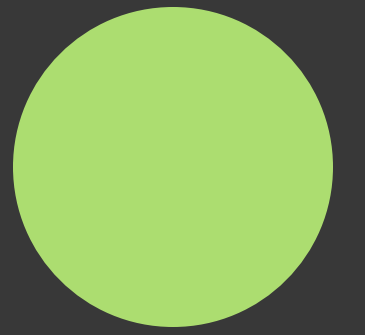
Tune intervention thresholds to prevent alert fatigue, providing help without interrupting the learner's flow.

03

Sensor Inaccuracy

Managing inconsistent hardware across multiple devices, as maintaining stable connections between the webcam, eye tracker, and PPG sensor was challenging.

Scaling & Roadmap



01

Autonomous Agentic Workflows

Move from reactive prompts to Agentic AI that can plan lessons, track goals, and adapt to each learner.

02

Cloud-Native Architecture

Offload inference to the cloud, making the system lightweight and browser-based.

03

Cost-Efficient Hardware Optimization

Optimize the vision pipeline for cheap webcams, eliminating costly eye-tracking hardware.

Applications



01

Adaptive Learning

Adjusts pacing and explanations in real time

02

Personalized Tutoring

Provides supportive, context-aware interventions

03

Cross-modal Tools

Adapt feedback across text, image, audio, and video

04

Educator Dashboards

Visualize learner engagement and challenges

Impact



01

Efficient Learning

Maintains optimal engagement, reducing fatigue and dropout

02

Proactive AI Partners

Anticipates learner needs and responds proactively

03

Adaptive Innovation

Enables cognition-aware and immersive learning systems

Thank You!