

The background of the slide is a photograph of a person with dark hair and glasses, seen from the back, sitting at a desk with two computer monitors. The image is covered with a semi-transparent blue overlay. The person is looking at the monitors, which display some graphical information. The overall tone is professional and tech-oriented.

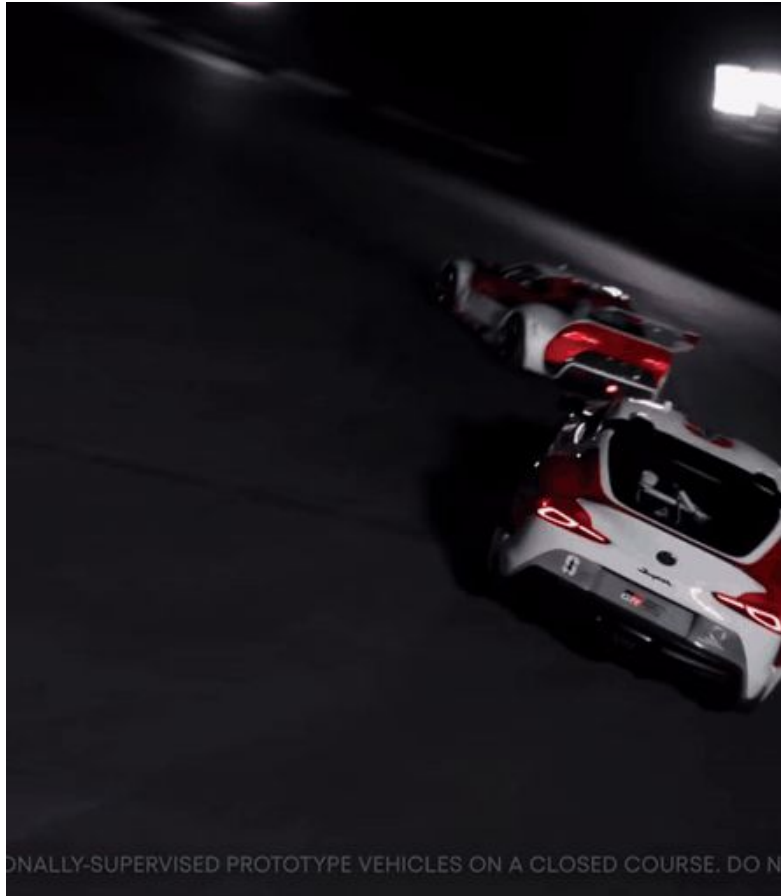
AI Coaching

Accelerating Human Skill Development with Reinforcement Learning

Wei Wang, Enlin Gu, Antonio Loquercio,
Haimin Hu, Rahul Mangharam

University of Pennsylvania / Johns Hopkins University

AI is becoming highly capable



Autonomous drifting



Drone racing

Can AI help us learn?

Human skills still matter

Physical control skills



Driving, drone racing
Teleoperation

Sports skills



Tennis
Badminton, hockey

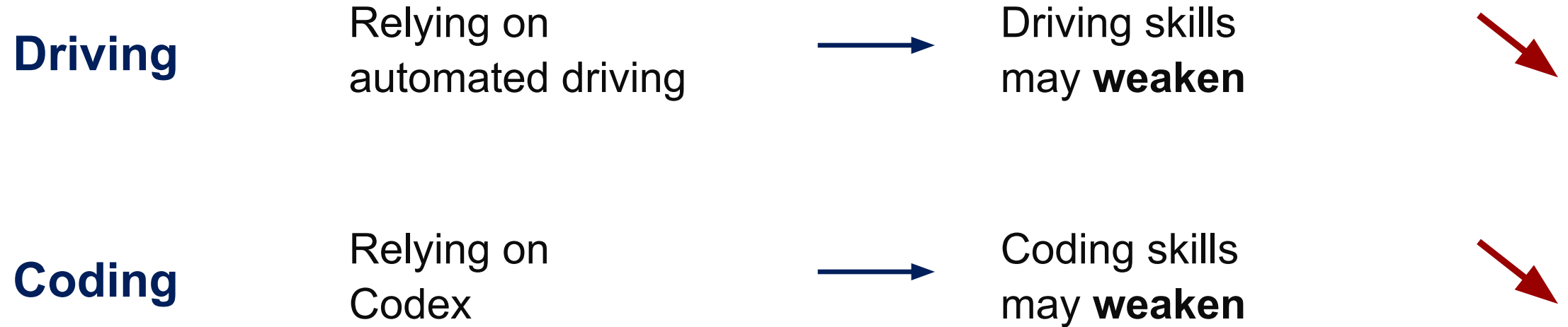
Human coaching is expensive

Core Question

**How can AI accelerate
human skill development?**

AI Assistance v.s. AI Coaching

Skill Development: AI assistance?



**Better assisted performance
is not independent performance**

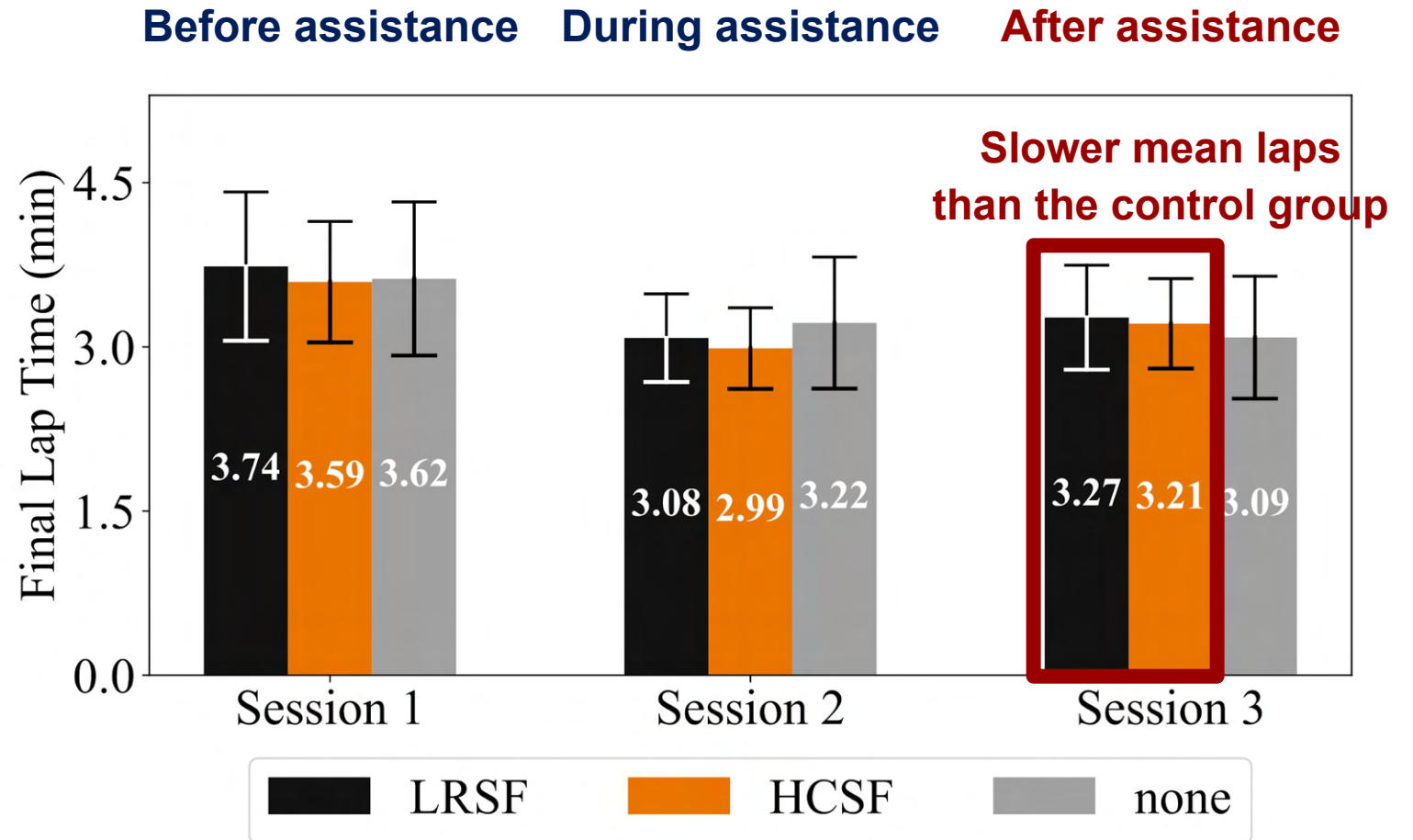
After assistance, mean performance is worse than in the unassisted group

Driving simulator

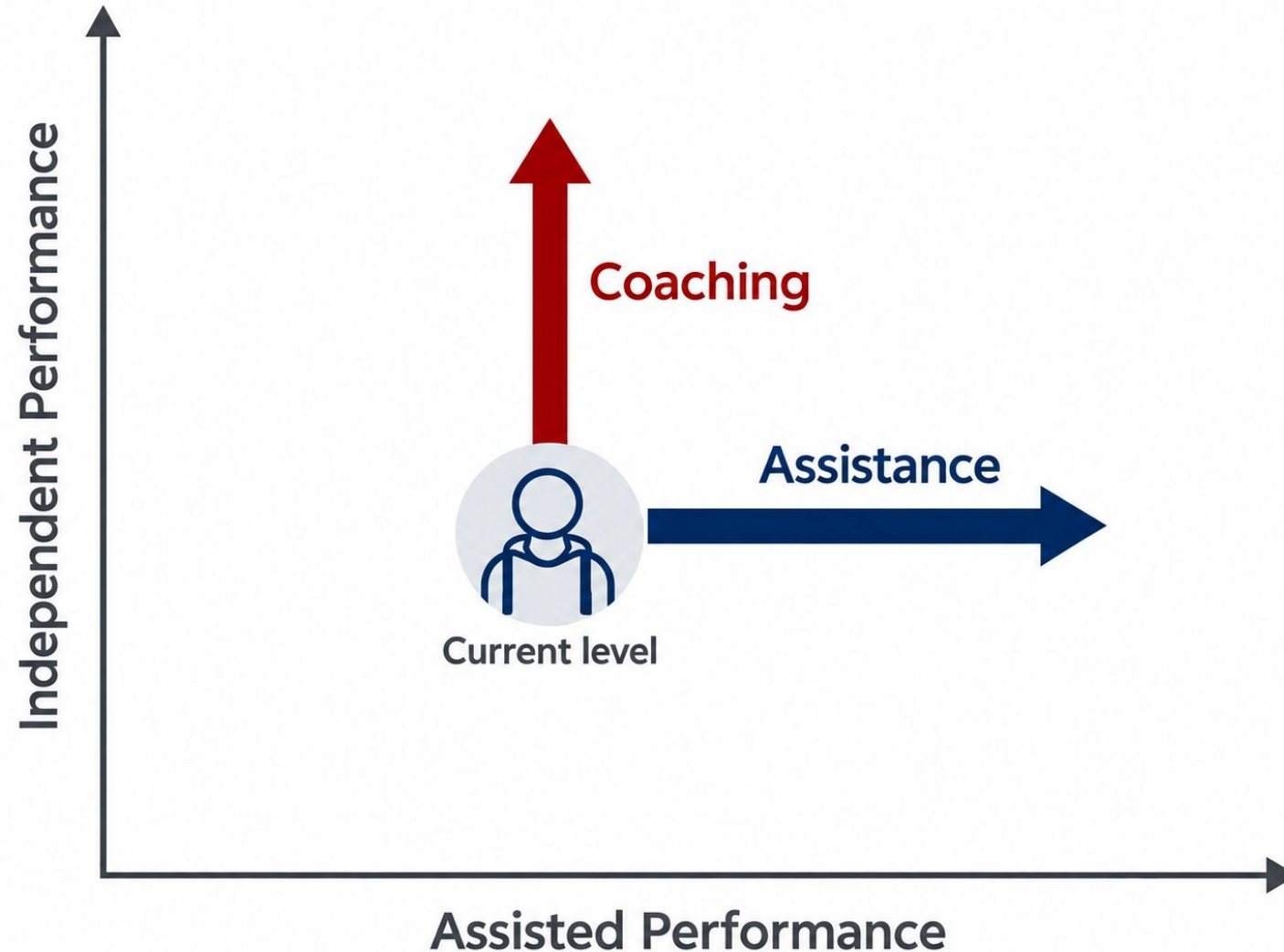


83 participants
Lower lap time = better

Gray: no-assistance control
in all three sessions



Goal of Assistance v.s. Coaching

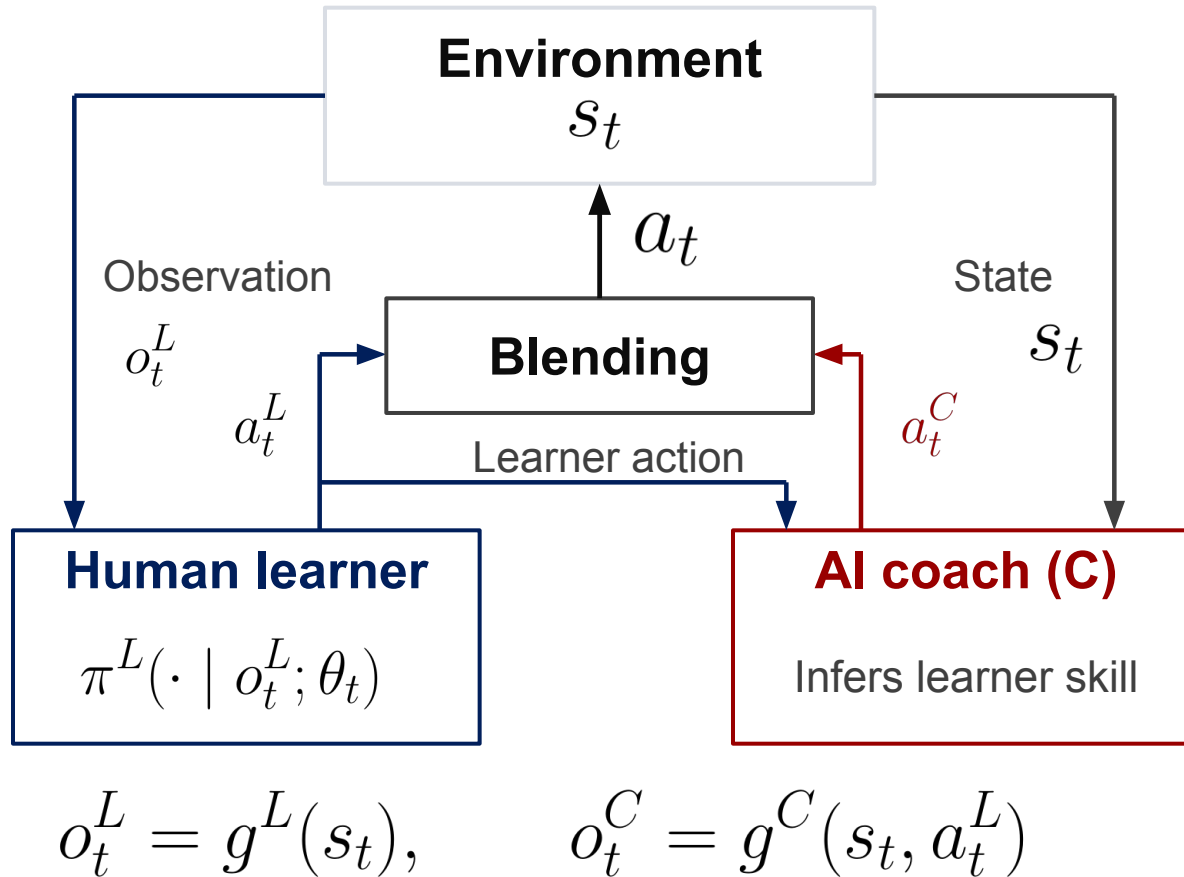


Problem Formulation

**How do we formulate
AI coaching?**

General-sum dynamic game

Two-player partially observable stochastic game



Learner goal

Assisted performance

$$r^L(s_t, a_t^L, a_t^C) = r^{\text{task}}(s_t, a_t)$$

Coach goal

Independent performance

$$r^C(s_t; \theta_t) = V^{\text{ind}}(s_t, \theta_t)$$

Distinct objectives, coupled through interaction

Simulated learner

Learner observation

$$o^L = g^L(s)$$

Physical state + task context

State $s, \theta \in [0, 1]$

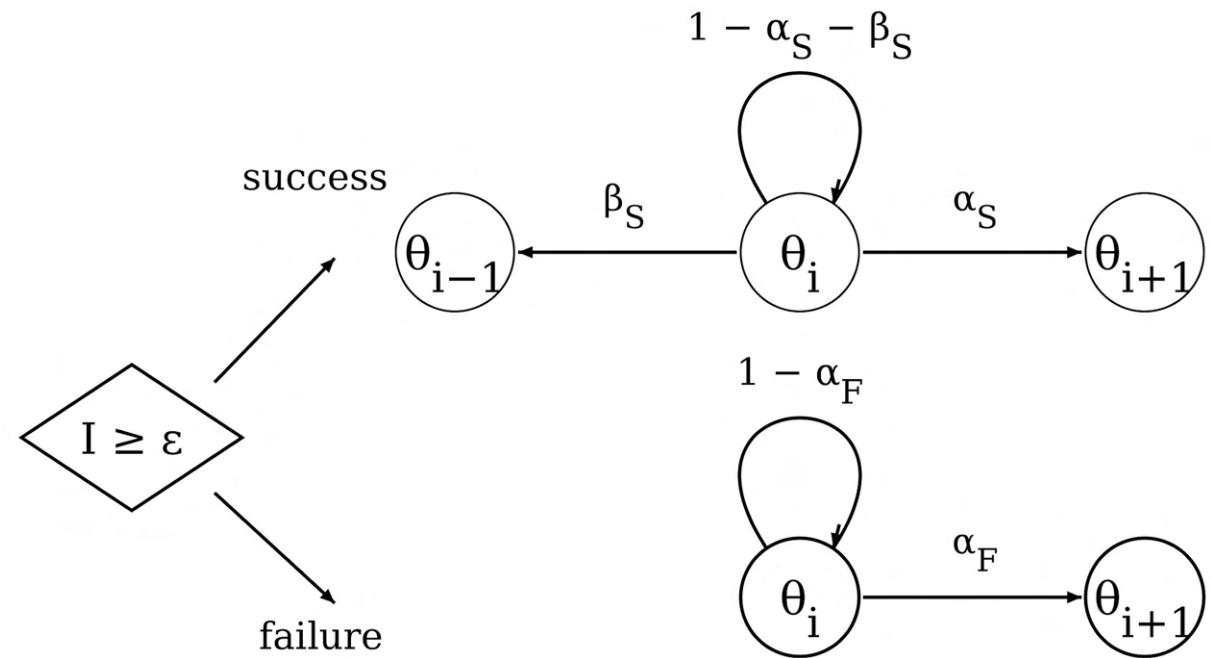
Reward $r^L = r^{\text{task}}(s, a)$

Boltzmann action model

$$\begin{aligned} \pi^L(a^L \mid o^L; \theta) \\ \propto \exp[\beta(\theta) Q^*(o^L, a^L)] \end{aligned}$$

Higher skill concentrates actions near the expert action.

Skill evolution



Small wins help novices;
skilled learners can learn from failure.

The simulated learner is part of the coach's training environment.

AI coach

Observation

Game formulation

$$o^C = g^C(s, \mathbf{a}^L)$$

Physical context + **learner action**

Current learner action, before blending

Training: known skill θ | Deployment: inferred skill $\hat{\theta}$

Output

$$\lambda = \pi_{\lambda}^C(o^C) \quad \text{per-axis assistance}$$

0 = learner; 1 = expert

$$a = \lambda \odot a^* + (\mathbf{1} - \lambda) \odot a^L$$

Blend first, then step
the environment

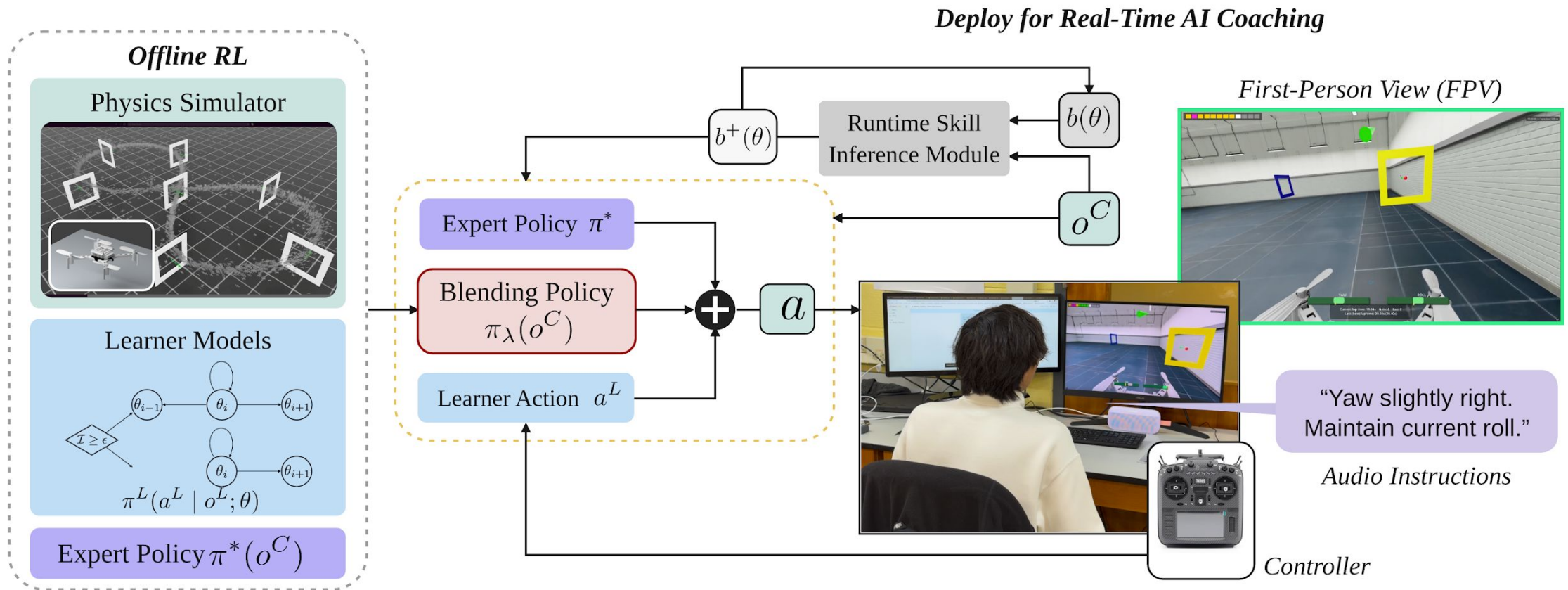
Reward Independent performance (Vol)

Expected task return without AI assistance

Training surrogate: $r_t^C = \theta_t - \theta_{t-1}$

Method

Learning to Coach (L2C)



6. Infer skill from behavior, then feed the estimate back to adapt assistance.

Adaptive shared control

Learn when to scaffold and when to step back

$$a_t = \lambda_t \odot \pi^*(o_t^C) + (\mathbf{1} - \lambda_t) \odot a_t^L$$

$$\lambda_t = \pi_\lambda^C(o_t^C) \in [0, 1]^{\dim(\mathcal{A})}$$

Per-axis assistance: 0 = learner, 1 = expert

Training in simulation

PPO uses simulated skill

$$r_t^C = \theta_t - \theta_{t-1}$$

Deployment with humans

Estimate skill from behavior

$$b^+(\theta) \propto p(o^C \mid \theta) b(\theta)$$

Experiments

Evaluation protocol

Independent performance, before and after coaching



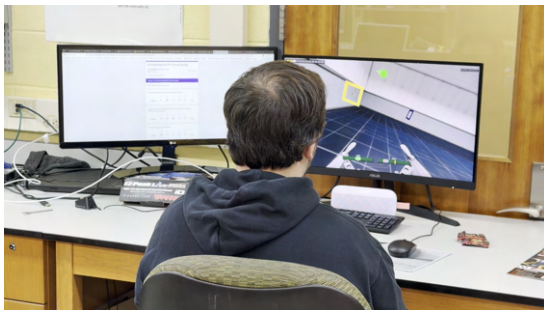
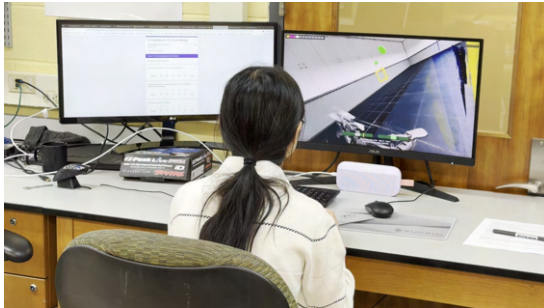
Lap time and failures per lap

Compare independent performance gains

User study: 33 participants

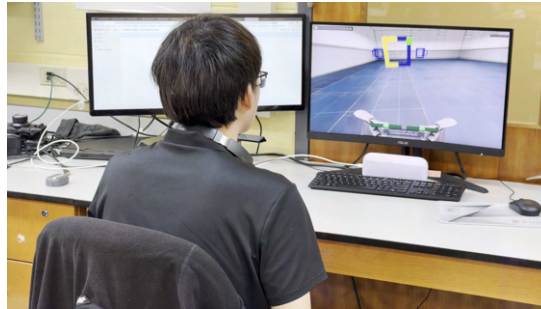
MIA · n = 11

Minimal intervention



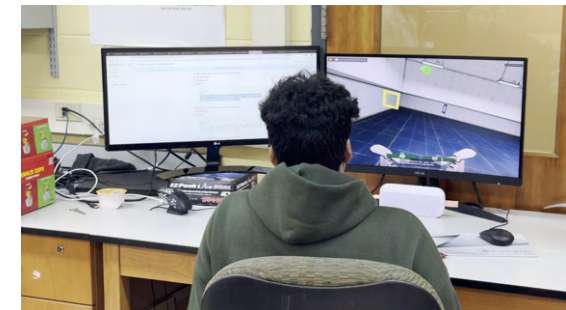
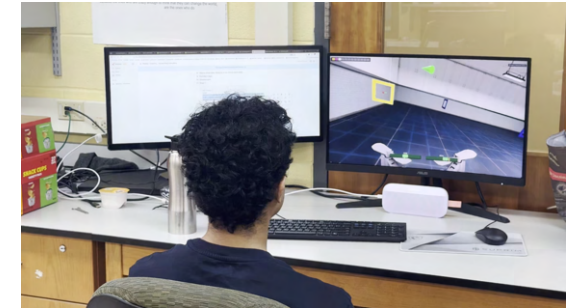
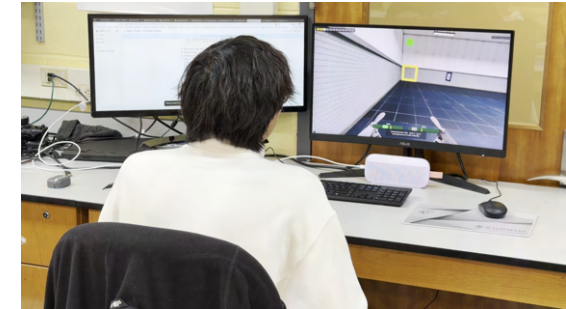
L2C · n = 11

Learning to Coach



RBF · n = 11

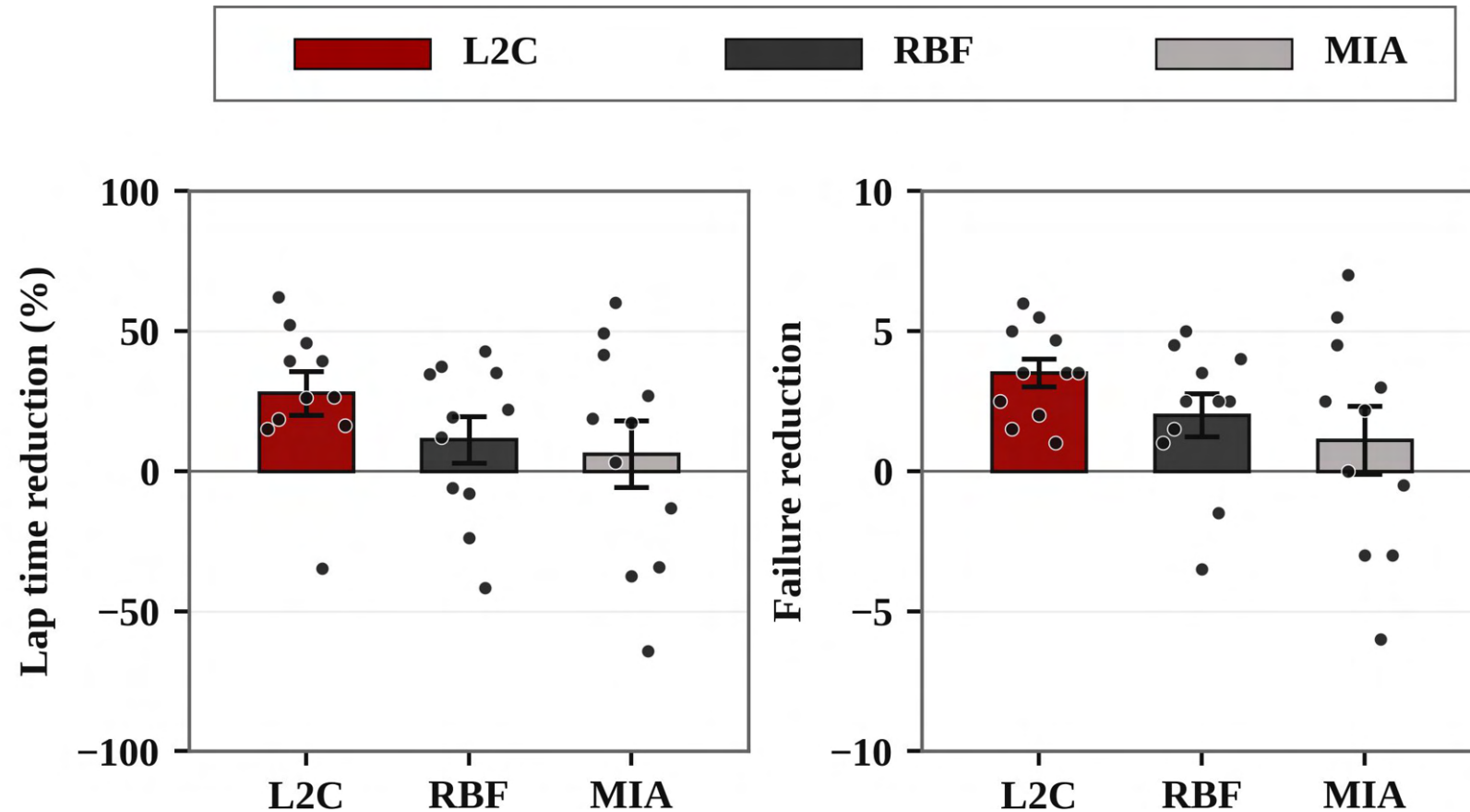
Fixed fading



Selected participants shown · randomly assigned to three methods

Independent performance

L2C: 27.9% shorter lap times; 3.52 fewer failures per lap

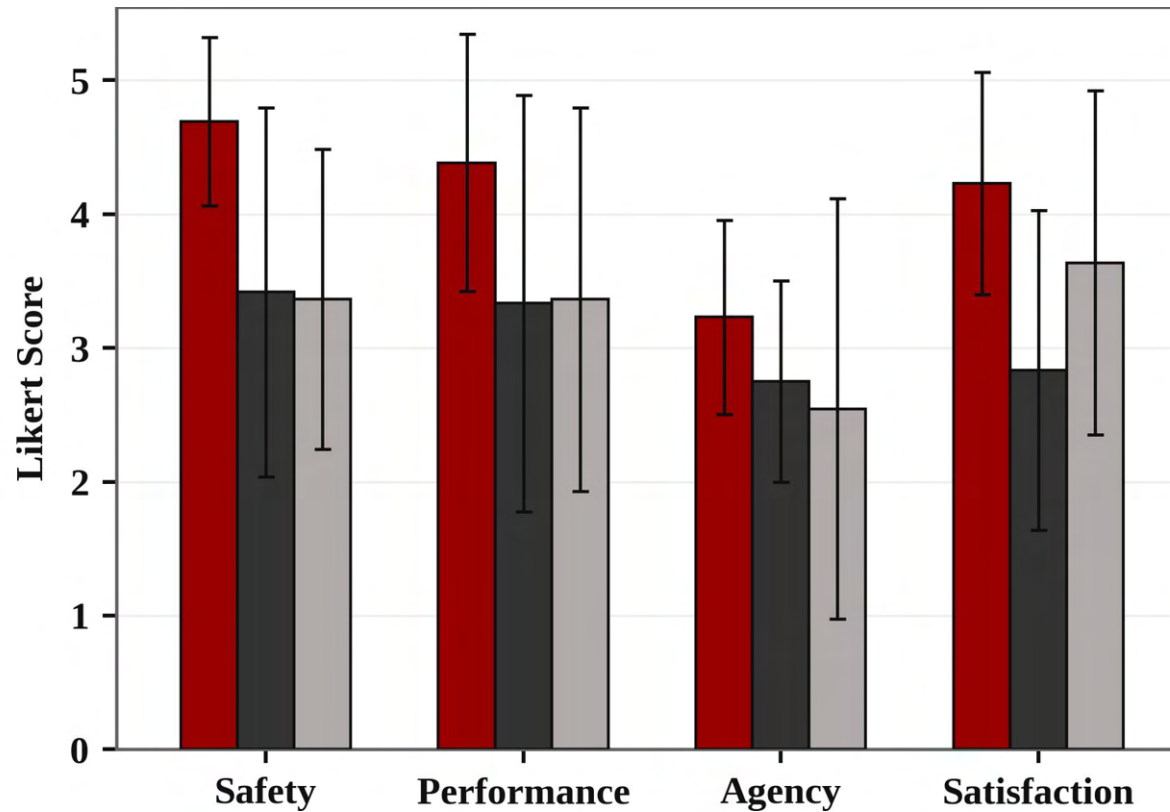


L2C pre/post gains significant; between-group tests not significant

Participant ratings

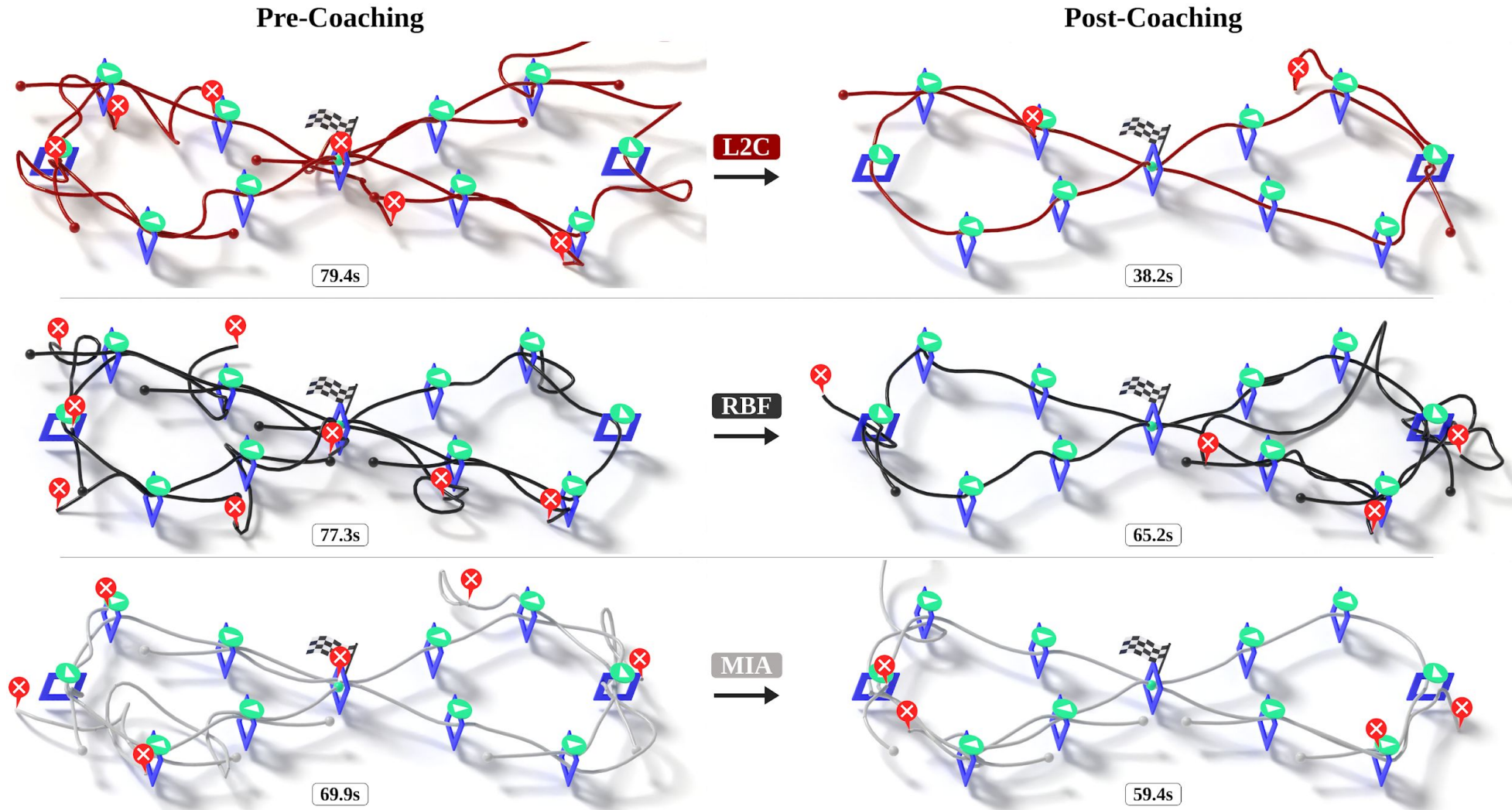
L2C received higher mean ratings in all four categories

■ L2C: Learning to Coach ■ RBF: fixed fading ■ MIA: minimal intervention



Descriptive ratings; independent performance is measured by the tests

Trajectories before and after coaching



AI coaching for human skill development

Improve independent performance

[Paper](#) · [Video](#)

Model coaching as a dynamic game

Learn assistance that adapts to skill

Evaluate independent performance



ai-coaching-drone-racing.github.io

Presentation TODO: Formulation

Post-presentation feedback · 15 Sep 2026 | Background: keep the current story

01 Give the audience a roadmap

Introduce ~3 modules before the details: Learner, Coach, and Environment / interaction.
Reuse module icons and highlight the current module on each page.

02 Explain the simulated learner

Explain how the learner acts, learns, and participates in coach training.
Define θ and other notation as introduced; distinguish skill θ from assistance λ .

03 Introduce skill estimation earlier

Motivate the belief model before the Method section: human skill is latent.
Distinguish known simulated skill from skill estimated from human behavior.

Presentation TODO: Method & Results

Revision checklist · not part of the talk | Keep slides and speaker notes in English

04 Correct the method diagram · slide 13

Change the leftmost label from “Offline Reinforcement Learning” to “Reinforcement Learning”; update matching reveal frames and notes.

05 Remove repeated explanation · slide 14

Remove or merge “Adaptive shared control” from the main talk.
Keep any essential definitions in slide 13 or the relevant formulation module.

06 Add the second qualitative result

Keep the trajectory figure and add the λ / assistance-level evolution figure.
Use the original paper / Keynote figure; verify axes, legends, and variable definitions.