



US DOT National
University Transportation Center for Safety

Carnegie Mellon University



Guardian Angel: Building Trust in Shared Autonomy Vehicles with Safe Human-Machine Interaction

Learning to Coach: AI that Builds Human Skill
Instead of Replacing It

Project #586

Rahul Mangharam, Principal Investigator

<https://orcid.org/0000-0002-3388-8283>

University of Pennsylvania, School of Engineering and Applied Science
200 South 33rd Street, Moore Building, Philadelphia, PA 19104

Final Report – October 2026

Federal Grant #69A3552344811

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, under grant number 69A3552344811 from the U.S. Department of Transportation's University Transportation Centers Program. The U.S. Government assumes no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No. Safety21 Project 586	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Guardian Angel: Building Trust in Shared Autonomy Vehicles with Safe Human-Machine Interaction		5. Report Date October 2026
7. Author(s) Rahul Mangharam, https://orcid.org/0000-0002-3388-8283		6. Performing Organization Code
9. Performing Organization Name and Address University of Pennsylvania, School of Engineering and Applied Science 200 South 33rd Street, Moore Building, Philadelphia, PA 19104		8. Performing Organization Report No. Safety21 Project 586
		10. Work Unit No.
		11. Contract or Grant No. Federal Grant #69A3552344811

12. Sponsoring Agency Name and Address U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program (UTC grant, Bipartisan Infrastructure Law) 1200 New Jersey Avenue SE, Washington, DC 20590 Safety21 National University Transportation Center, Carnegie Mellon University (UTC grant recipient) 5000 Forbes Avenue, Pittsburgh, PA 15213		13. Type of Report and Period Covered Final Report (July 1, 2025 – June 30, 2026)
		14. Sponsoring Agency Code USDOT
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation and the Safety21 University Transportation Center. Federal share: \$35,000 in UTC funds. Non-federal match: \$35,000 University of Pennsylvania cost share (1:1). Project record: https://ppms.cit.cmu.edu/projects/detail/586 . Project website: https://ai-coaching-drone-racing.github.io/ . Publication identifier: https://doi.org/10.48550/arXiv.2606.25337 . Full list of research products in Section 8.		
16. Abstract Project 586 studies how shared autonomy can build trust and safety in automated vehicles through safe human-machine interaction. In shared autonomy, an automated agent supervises a person who keeps control, acting as a guardian angel. A concern raised by recent studies is that constant assistance can leave drivers over-reliant and less skilled when the assistance is absent. The project therefore reframes the guardian angel as a coach. It formalizes coaching as a game in which the learner optimizes task performance while the coach targets the learner's independent competence, and it trains a coaching policy with reinforcement learning that decides when to help and when to step back. The method, Learning to Coach (L2C), was tested with 33 participants in a first-person-view drone racing simulator and compared with two state-of-the-art baselines. After about 40 minutes of coaching, L2C participants cut lap time by 27.9 percent and made 3.52 fewer failures per lap, and L2C was the only method with significant within-participant gains on both measures. Participants rated L2C higher on safety, performance, agency, and satisfaction. The work is accepted to the Conference on Robot Learning (CoRL) 2026.		
17. Key Words Shared autonomy; AI coaching; human-machine interaction; reinforcement learning; driver trust; skill development; drone racing; user study		18. Distribution Statement No restrictions. This document is available through the Safety21 project record.
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 17
		22. Price

Contents

1	Project Overview	1
1.1	Project Funding and Matching Funds	1
1.2	Summary of Results	2
2	Why Shared Autonomy Must Protect Human Skill	2
3	Learning to Coach	3
3.1	The Human-AI Coaching Game	3
3.2	Reinforcement Learning Framework	3
3.3	Implementation	4
4	User Study and Results	4
4.1	Protocol	4
4.2	Results	5
4.3	Limitations	6
5	Status Against the Project Proposal	6
6	Collaboration, Impact, and Workforce	7
6.1	Collaboration	7
6.2	Impact on the Transportation System	7
6.3	Workforce Development	8
6.4	Dissemination	8
7	Next Steps and Recommendations	8
8	Research Products, Data, and Dissemination	8
8.1	Publications	8
8.2	Software, Data, and Websites	8
8.3	Datasets and Metadata	8
8.4	Research Personnel and ORCID Information	9
A	Abbreviations	10

List of Figures

1	The Learning to Coach system. The coaching policy is trained in simulation with a learner model and an expert policy, then deployed to blend assistance and give feedback to a human learner. Figure from [1].	4
2	Example flights before and after coaching with each method. Crosses mark failures and the flag marks the start and finish. Figure from [1].	5
3	After coaching, participants trained with L2C show larger reductions in lap time and failure count (left) and give higher ratings for safety, performance, agency, and satisfaction (right). Figure from [1].	5
4	The L2C coach assigns more assistance to a novice (left, estimated skill 0) than to a more skilled learner (right, estimated skill 0.5), and varies assistance around the gates. Figure from [1].	6

List of Tables

1	Project 586 funding sources.	2
2	Main quantitative results of the user study.	2
3	Proposal items and their status at the end of the project period.	7
4	Publication supported by the project.	9
5	Websites and supporting material.	9
6	Principal investigator and ORCID information.	9
7	Abbreviations used in this report.	10

1 Project Overview

Safety21 Project 586, Guardian Angel: Building Trust in Shared Autonomy Vehicles with Safe Human-Machine Interaction, studies how people and automated driving agents can share control of a vehicle safely and with justified trust. In shared autonomy, a person keeps control while an automated agent supervises and corrects errors. The agent acts as a guardian angel. This arrangement offers a path to wider use of automation, because the person can intervene when the automation is uncertain, and the automation can step in when the person errs.

The project began from the proposal to expose drivers to manual, fully automated, and guardian-angel driving in an immersive simulator and to measure comfort and trust. As the research developed, the team identified a question that sits underneath trust: what happens to a person's own skill when an automated guardian is always present? Recent studies indicate that persistent assistance can leave people over-reliant and less skilled when the assistance is removed. A safe shared-autonomy system should therefore support the person's independent competence, and not only prevent the next mistake. The Safety21 sponsor allows projects to depart from the original proposal. Section 5 compares the work with the proposal in detail.

Led by Principal Investigator Rahul Mangharam at the University of Pennsylvania, the project produced one main result.

1. **Learning to Coach (L2C).** A reinforcement learning framework that trains an AI coach to decide when to help a human learner and when to step back [1]. It was tested in a user study with 33 participants flying a first-person-view (FPV) drone racing simulator, and it is accepted to the Conference on Robot Learning (CoRL) 2026.

Section 2 explains why shared autonomy needs to protect human skill. Section 3 describes the method, and Section 4 describes the user study and its results. Section 5 compares the outputs with the proposal. Section 6 covers collaboration, impact, and workforce, Section 7 gives the next steps, and Section 8 lists the research products.

Overall project outcome

Project 586 delivered a method that turns an automated guardian into a coach and showed, with 33 participants, that it speeds up human learning. After about 40 minutes of coaching, participants who used the L2C coach cut lap time by 27.9 percent and made 3.52 fewer failures per lap when flying alone, and L2C was the only method with significant gains on both measures. Participants also rated the L2C coach higher on safety, performance, agency, and satisfaction. The result gives shared-autonomy designers a principled way to protect and build human skill. Validation so far is in a simulator, and on-road driving studies remain future work.

1.1 Project Funding and Matching Funds

Project 586 was funded through the Safety21 National University Transportation Center under U.S. Department of Transportation grant 69A3552344811. Table 1 states the source of all project funds. The University of Pennsylvania provided the full non-federal match, equal to the federal award, as institutional cost share. No other matching sources were used. The federal share was fully expended by the end of the project period.

Table 1: Project 586 funding sources.

Funding source	Type	Amount
U.S. Department of Transportation, UTC Program, through Safety21 (grant 69A3552344811)	Federal	\$35,000
University of Pennsylvania institutional cost share	Non-federal match	\$35,000
Total project budget		\$70,000

1.2 Summary of Results

Table 2 lists the main quantitative results.

Table 2: Main quantitative results of the user study.

Result	Evidence	Section
Faster independent performance	L2C participants reduced lap time by 27.9 percent on average after about 40 minutes of coaching (paired t -test $p = 0.005$, Cohen’s $d_z = -1.08$).	4.2
Fewer failures	L2C participants made 3.52 fewer failures per lap ($p < 0.001$, $d_z = -2.13$). Failures are collisions and timeouts.	4.2
Baselines	Rule-based fading reduced lap time by 11.3 percent ($p = 0.21$) and failures by 2.0 per lap ($p = 0.027$). Minimally invasive assistance reduced lap time by 6.2 percent ($p = 0.62$) and failures by 1.11 per lap ($p = 0.38$).	4.2
Participant experience	L2C received higher mean ratings than both baselines on safety, performance, agency, and satisfaction.	4.2
Scale of study	33 participants, randomly assigned to three coaches (11 each).	4.1

2 Why Shared Autonomy Must Protect Human Skill

Fully automated driving is not yet able to handle every situation. Sensors and algorithms struggle with unusual weather, erratic road users, and work zones. Shared autonomy keeps a human in the loop for these cases and lets the automated agent monitor and correct the human. This blend of human judgment and machine precision is widely seen as a practical route to safer roads while automation matures. It also supports trust, because people who know they can take control are more willing to use the technology.

Shared control has been demonstrated for drone piloting and high-speed car racing, where an AI copilot fuses the human command with its own corrective input [3, 4, 2]. These systems improve the joint performance of the human and the machine. The same studies point to a cost. A study of 83 participants in a driving simulator found that after a period of assistance, participants who had been assisted flew slower mean laps than a control group that never received assistance [2]. Better assisted performance is not the same as better independent performance. Broader work on AI assistance describes the same risk as over-reliance and skill stagnation, because the person is denied the productive failures that drive learning.

This risk matters for transportation. A driver who relies on a guardian every day may not be prepared for the moment when the guardian is absent, degraded, or has handed control back. Trust that is built on an assistant that hides the driver’s own weaknesses is not well calibrated. The project therefore asks how a guardian angel

can keep the person safe and also leave the person more capable.

- **When should the agent step in?** Intervening too much deprives the learner of useful failures, and intervening too little exposes the learner to failures that are uninformative or discouraging.
- **How can the agent tell?** The learner’s skill is not directly observable, so the agent must infer it from behavior and adapt.
- **How should success be measured?** The test is the learner’s performance without the assistant, not performance with it.

These questions connect to the priorities of the U.S. Department of Transportation for automated driving systems (ADS), advanced driver-assistance systems (ADAS), artificial intelligence and machine learning, and robotics.

3 Learning to Coach

3.1 The Human-AI Coaching Game

The project formalizes interactive coaching as a two-player partially observable stochastic game between a human learner and an AI coach. The two players share control of the vehicle but have different goals. The learner optimizes task performance with assistance. The coach is rewarded for the learner’s independent competence, which is the expected task return the learner would obtain if the coach withdrew assistance. This reward is called the value of independence. Because the players’ goals differ, the game is non-cooperative. Maximizing it gives the coach an incentive to scaffold when the learner needs help and to step back when a mistake can teach more than a rescue.

3.2 Reinforcement Learning Framework

Directly solving the game is intractable, because the learner’s skill is hidden and the coach’s influence on skill is hard to model. The project makes the problem tractable with three components, shown in Figure 1.

1. **Adaptive shared control.** The executed action is a blend of an expert policy and the learner’s command, with a per-axis blending level between zero (learner in control) and one (expert in control). The coach policy chooses the blending level from the physical context, and the blend is applied in closed loop, so assistance can rise near a gate and fall elsewhere.
2. **A skill-evolution model.** A probabilistic finite-state automaton describes how a learner’s skill moves up or down after successes and failures. It captures the causal influence of coaching on skill, and it is used with a skill-conditioned model of learner behavior to create simulated learners for training.
3. **A tractable training reward.** The coach is trained with proximal policy optimization (PPO) [7] on a surrogate reward equal to the change in skill. The paper proves that a coach which raises the learner’s probability of improving also earns a higher value of independence.

At deployment the learner’s skill is unobserved. The coach keeps a belief over the skill at each gate and updates it by Bayesian inference from the time the learner takes to pass the gate. The coach also gives verbal

instructions and visual cues, and after a failure it resets the learner to one gate before the failed segment for guided practice.

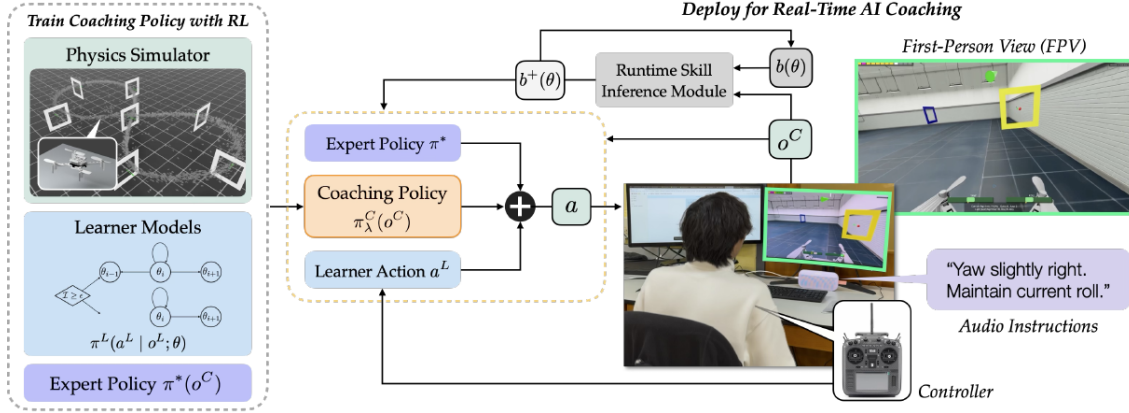


Figure 1: The Learning to Coach system. The coaching policy is trained in simulation with a learner model and an expert policy, then deployed to blend assistance and give feedback to a human learner. Figure from [1].

3.3 Implementation

The simulator is built in Isaac Lab on top of an existing drone racing environment. The track has 12 gates in a figure-eight layout. The expert policy and the coaching policy were each trained with PPO on a single NVIDIA RTX 5080 GPU, in about 8 hours for the expert and 6.5 hours for the coach. The human controls yaw and roll rate with a handheld remote controller, and the system handles pitch rate and thrust so that a novice can learn to steer within the session. The system logs the human, expert, and blended actions, the drone state, gate events, failures, lap times, and assistance level during every trial.

4 User Study and Results

4.1 Protocol

Participants were randomly assigned to one of three coaches, with 11 participants in each group. The coaches were L2C, rule-based fading (RBF), and minimally invasive assistance (MIA). RBF gradually fades assistance over the session on a fixed curve [5]. MIA is a reinforcement-learned copilot that changes the learner’s action as little as needed to stay safe and complete the task, in the style of human-centered safety filters [2]. All three coaches used the same simulator, track, interface, verbal instructions, and visual cues, so the only difference was the assistance method. Each participant completed three phases.

1. A pre-coaching test of two laps flown alone.
2. A coached session of 15 laps, about 40 minutes. After every three coached laps the participant flew one lap alone, which gave the coach a clean estimate of skill.
3. A post-coaching test of two laps flown alone.

Skill change was measured by lap time and by failure count, where a failure is a collision or a gate not passed

within 8 seconds. After the trial, participants answered a five-point questionnaire on perceived safety, performance, agency, and satisfaction. The study population was mostly novices. Of the participants, 61.1 percent reported no drone-flying experience, 36.1 percent reported casual experience, and 2.8 percent reported regular experience. A statistical test found no significant difference in initial self-reported skill among the three groups.

4.2 Results

Table 2 summarizes the within-participant results. L2C was the only coach with significant improvement in both lap time and failure count. For RBF, lap time did not change reliably and the failure reduction was about one-third the size of the L2C effect. For MIA, neither measure changed reliably. Figure 2 shows example flights before and after coaching, and Figure 3 shows the group results and participant ratings.

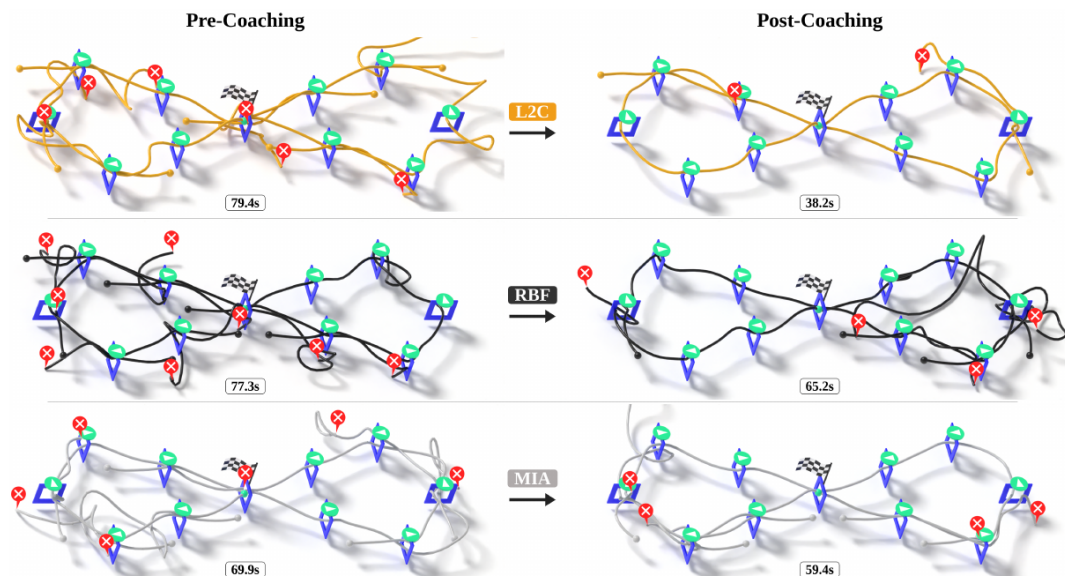


Figure 2: Example flights before and after coaching with each method. Crosses mark failures and the flag marks the start and finish. Figure from [1].

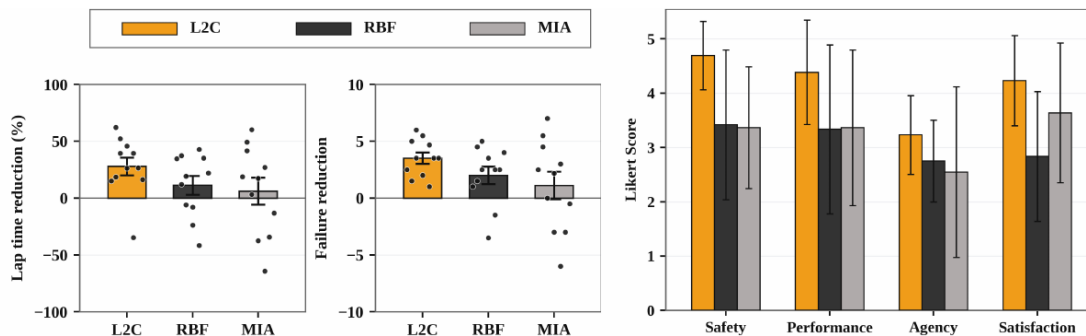


Figure 3: After coaching, participants trained with L2C show larger reductions in lap time and failure count (left) and give higher ratings for safety, performance, agency, and satisfaction (right). Figure from [1].

Welch tests compared L2C with each baseline, with Holm correction. L2C outperformed MIA by 21.7 per-

centage points in lap time reduction and by 2.41 failures per lap, and it outperformed RBF by 16.6 percentage points and 1.52 failures per lap. All four contrasts favored L2C with medium to large effect sizes (Hedges' g between 0.60 and 0.76). With 11 participants per group these between-method differences did not reach conventional statistical significance (p between 0.09 and 0.16), and the paper reports them as supporting evidence, not as proof. The result that L2C alone produced significant within-participant gains on both outcomes is the main finding.

Figure 4 shows how the coach adapts. A less skilled learner (estimated skill 0) receives more assistance than a more skilled learner (estimated skill 0.5). Within a lap the assistance level changes sharply near the gates, which are the task-critical moments that decide whether the learner collides or passes.

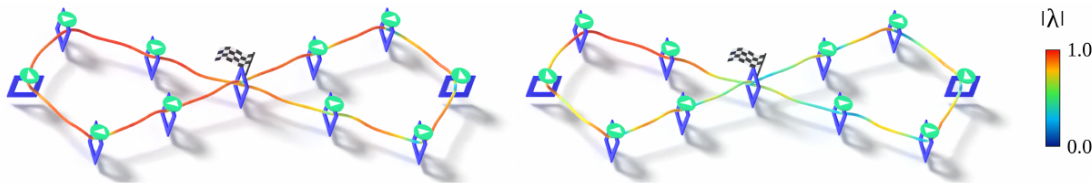


Figure 4: The L2C coach assigns more assistance to a novice (left, estimated skill 0) than to a more skilled learner (right, estimated skill 0.5), and varies assistance around the gates. Figure from [1].

4.3 Limitations

The study used simulated drone racing with a restricted two-axis control interface, and the learner model used in training may not capture the full variability of real human learning. Verbal and visual cues were predefined and triggered by fixed rules, and only the physical assistance level was learned. The participants were mostly novices, and the sample was small. These limits define the next steps in Section 7.

5 Status Against the Project Proposal

Table 3 compares the outputs and milestones in the proposal with the work completed. The project changed the experimental setting and the participant pool as the research developed. The main change is that the study was carried out in an FPV drone racing simulator with 33 mostly novice participants, instead of a driving simulator with 20 drivers in two age groups. The Safety21 sponsor allows such deviation. The change let the project test the central question of the proposal, how a guardian agent shapes human behavior and trust, in a setting where learning can be measured in a single session.

Table 3: Proposal items and their status at the end of the project period.

Proposal item	Status	Evidence or plan
Guardian Angel technique that lets a person drive while an automated agent supports them	Complete, in a new form	Adaptive shared control with an AI coach that adjusts assistance to the learner’s skill (Section 3).
Immersive simulator and driver interface	Complete, in a new setting	Isaac Lab FPV drone racing simulator with a handheld controller, first-person view, audio, and visual cues.
User study with surveys and simulator data	Complete, with a different cohort	33 participants, three coaches, trial logs, and a post-study questionnaire (Section 4).
Cohorts of young and mature drivers, 20 drivers in total	Changed	33 participants, mostly novices to drone piloting. Age-group comparison is not part of the result.
Driving scenario in the MetaDrive simulator and Autoware connection	Not carried out in this period	Carrying the coach to driving simulators and Autoware is a next step (Section 7).
Publications at major venues	Complete	One paper accepted to CoRL 2026.
Source code and software	Not yet public	Project website is public. Release of code and data is planned.
Training of students	Supported	Penn researchers carried out the research reported in the paper (Section 6.3).

6 Collaboration, Impact, and Workforce

6.1 Collaboration

The research was carried out at the University of Pennsylvania in collaboration with Haimin Hu at Johns Hopkins University, who is co-advisor of the work and co-author of the paper. The simulator builds on an existing agile-flight racing environment [6]. No formal partnership with an industry or agency partner was part of the work in this period.

6.2 Impact on the Transportation System

The results bear on how automated driving should be designed and regulated.

- **Safety.** A guardian that also builds skill helps drivers who must take over when automation fails or is unavailable. The study showed lower failure counts when flying alone after coaching.
- **Trust.** Participants rated the adaptive coach higher on perceived safety, performance, agency, and satisfaction. Preserving a sense of agency is a condition for calibrated trust.
- **Policy and design.** The framework gives designers and regulators a measurable criterion for human-centered assistance, namely independent performance after assistance, in addition to performance during assistance.
- **Training.** The same method applies to driver training, teleoperation, and other motor skills where human skill and automation share a task.

6.3 Workforce Development

The paper was written by Wei Wang, Enlin Gu, Antonio Loquercio, and Rahul Mangharam of the University of Pennsylvania and Haimin Hu of Johns Hopkins University and Penn. The work trained researchers in reinforcement learning, game-theoretic modeling, GPU-accelerated simulation, human-subject experiment design, and statistical analysis of user studies. These skills are in demand for human-centered automated driving and robotics.

6.4 Dissemination

The paper is accepted to CoRL 2026 and is available as an arXiv preprint [1]. A public project website at <https://ai-coaching-drone-racing.github.io/> gives the paper and video.

7 Next Steps and Recommendations

- **Bring the coach to driving.** Apply L2C to car driving simulators and to shared-control driving with an automated driving stack such as Autoware, and measure how coaching affects takeover skill and trust.
- **Study experienced and diverse drivers.** Recruit larger groups, including drivers of different ages and experience, to test how benefits differ across learners, as in the original proposal.
- **Learn the full coaching policy.** Learn the verbal and visual feedback together with the physical assistance, and replace the hand-built learner model with learner models from human data.
- **Retention.** Measure whether gains persist days or weeks after coaching.
- **Release.** Publish the code and the anonymized user-study data with metadata.

Barriers. Validation so far is in simulation with a small sample. On-road studies will need institutional review of human-subject protocols, safety plans, and partners who operate test vehicles.

8 Research Products, Data, and Dissemination

The project produced one publication, listed in Table 4.

8.1 Publications

8.2 Software, Data, and Websites

8.3 Datasets and Metadata

The user study logged the human, expert, and blended actions, the drone state, gate passage events, failures, lap and segment times, and the assistance level for each trial, together with the questionnaire responses. The metadata describe the simulator and software versions (Isaac Sim 5.1.0, Isaac Lab 0.48.5), the track layout,

Table 4: Publication supported by the project.

Short name	Citation and links	Venue and status
L2C	W. Wang, E. Gu, A. Loquercio, H. Hu, and R. Mangharam, “AI Coaching for Accelerating Human Skill Development with Reinforcement Learning,” accepted to the <i>Conference on Robot Learning (CoRL)</i> , 2026. arXiv: https://arxiv.org/abs/2606.25337 Website: https://ai-coaching-drone-racing.github.io/	Conference paper, accepted

Table 5: Websites and supporting material.

Item	Purpose	Access
AI Coaching project website	Paper, video, and project description.	https://ai-coaching-drone-racing.github.io/

the coach and baseline settings, the study phases, the number of laps, and the metrics reported in the paper. The records are maintained by the authors under the project Data Management Plan (DMP). Public release of the anonymized data is planned (Section 7).

8.4 Research Personnel and ORCID Information

Table 6: Principal investigator and ORCID information.

Name	Project or publication role	ORCID
Rahul Mangharam	Principal investigator; author of the publication	https://orcid.org/0000-0002-3388-8283

External co-authors are affiliated with other institutions and are listed in the publication.

References

- [1] W. Wang, E. Gu, A. Loquercio, H. Hu, and R. Mangharam, “AI coaching for accelerating human skill development with reinforcement learning,” accepted to *Conference on Robot Learning*, 2026. arXiv:2606.25337. [Online]. Available: <https://arxiv.org/abs/2606.25337>
- [2] D. D. Oh, J. Lidard, H. Hu, H. Sinhmar, E. Lazarski, D. Gopinath, E. S. Sumner, J. A. DeCastro, G. Rosman, N. E. Leonard *et al.*, “Safety with agency: Human-centered safety filter with application to AI-assisted motorsports,” in *Proc. Robotics: Science and Systems*, 2025.
- [3] S. Reddy, A. D. Dragan, and S. Levine, “Shared autonomy via deep reinforcement learning,” in *Proc. Robotics: Science and Systems*, 2018.
- [4] J. DeCastro, A. Silva, D. Gopinath, E. Sumner, T. M. Balch, L. Dees, and G. Rosman, “Dreaming to assist: Learning to align with human objectives for shared control in high-speed racing,” in *Conference on Robot Learning*, 2024.

- [5] C. Shen, S. Yu, Y. Weng, H. Ma, C. Li, H. Yasuda, J. Dallas, M. Thompson, J. Subosits, and T. Ersal, “Cyber racing coach: A haptic shared control framework for teaching advanced driving skills,” arXiv:2509.20653, 2025.
- [6] V. Pasumarti, L. Bianchi, and A. Loquercio, “Agile flight emerges from multi-agent competitive racing,” arXiv:2512.11781, 2025.
- [7] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” arXiv:1707.06347, 2017.
- [8] Safety21 University Transportation Center, “Project 586: Guardian Angel: Building trust in shared autonomy vehicles with safe human-machine interaction.” Accessed: Oct. 6, 2026. [Online]. Available: <https://ppms.cit.cmu.edu/projects/detail/586>

A Abbreviations

Table 7: Abbreviations used in this report.

Abbreviation	Definition
ADAS	Advanced driver-assistance system
ADS	Automated driving system
CoRL	Conference on Robot Learning
DMP	Data Management Plan
FPV	First-person view
GPU	Graphics processing unit
L2C	Learning to Coach
MIA	Minimally invasive assistance
PPMS	Safety21 project and performance management system
PPO	Proximal policy optimization
RBF	Rule-based fading
USDOT	U.S. Department of Transportation
UTC	University Transportation Center
VoI	Value of independence