Proximal State Nudging: Reducing Skill Atrophy from AI Assistance

Anonymous Authors[1]

Abstract

The gradual loss of human skills, or skill atrophy, is a rising concern as users increasingly rely on AI assistance. This problem is particularly salient in cooperative AI systems, such as aircraft piloting or driving, where humans and AI agents jointly share control and decision-making. In these settings, human operators often struggle to disentangle which outcomes arise from AI intervention versus their own actions, undermining opportunities for their own learning and long-term skill retention. In this work, we propose PROXIMAL STATE NUDGING, a cooperative shared-control algorithm that balances assistance with human skill development. Rather than optimizing solely for combined Human+AI task reward, our method also gradually "nudges" human users toward states in the environment where they are most likely to improve their own, unassisted competence. In two simulated environments (Discrete MDP and LunarLander), Proximal State Nudging outperforms existing shared autonomy baselines in improving a student's unassisted performance. We further validate our approach through two human subject studies (Parallel Parking and High Performance Racing, $n = 60$) using the high-fidelity CARLA driving simulator, showing that we can build real-world cooperative AI systems that support human agency and skill retention without sacrificing performance.

1. Introduction

Skill atrophy resulting from AI assistance has been documented across domains including transportation, medicine, and education (de Winter et al., 2023a; Sarofim, 2024; Casner et al., 2016; Macnamara et al., 2024). As AI systems increasingly participate in high-stakes decision-making and control, there is growing concern that automation designed to optimize short-term performance may inadvertently erode human competence, agency, and resilience over time. This challenge is particularly important for cooperative AI systems where humans and AI agents must work together effectively while preserving human oversight and capability.

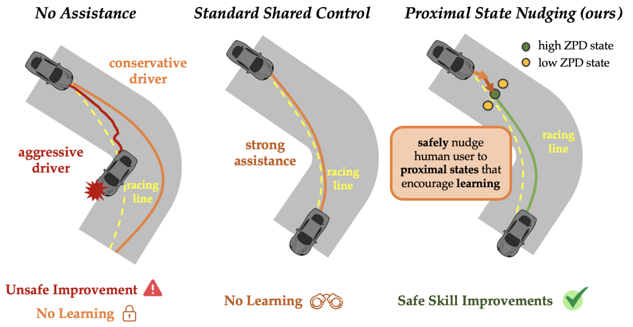


Figure 1: **Overview.** In a cooperative human+AI driving task, human users receiving *No Assistance* from AI may engage in unsafe exploration or remain in their comfort zone and fail to learn. In *Standard Shared Autonomy*, strong AI assistance can suppress exploration and lead to skill atrophy. We propose, PROXIMAL STATE NUDGING, which gently nudges users toward states estimated to fall within the user's *Zone of Proximal Development* (ZPD), maximizing their learning gains while still maintaining safety.

Consider domains such as driving or aircraft operation, where AI assistance can improve immediate safety and performance, yet human operators must still remain capable of intervening when automation fails or circumstances change (Endsley, 2017). Existing approaches to mitigating skill atrophy often frame assistance and learning as competing objectives: either maximize performance through automation or reduce assistance to encourage human learning (Dell'Aquila et al., 2023). However, AI systems designed for societal benefit should ideally support both immediate task success and long-term human development.

In this work, we introduce PROXIMAL STATE NUDGING, a learning-aware shared autonomy framework inspired by the Zone of Proximal Development from educational psychology (Vygotsky, 1978). Rather than optimizing solely for task completion, PROXIMAL STATE NUDGING guides users toward states where learning and skill acquisition are most likely to occur while maintaining strong overall task performance. By incorporating principles of cognitive scaffolding into cooperative AI, our approach promotes sustained human competence alongside effective human–AI collaboration.

Unlike prior learning-compatible assistance methods, PROXIMAL STATE NUDGING does not require access to the human learner's internal value function or an explicit model of how AI influences learning. We evaluate the approach in simulated environments (DiscreteMDP and LunarLander) as well as two human-subject driving studies (Parallel Parking and High Performance Racing, $n = 60$) in the high-fidelity CARLA driving simulator. Across settings, our results demonstrate that cooperative AI systems can mitigate skill atrophy while preserving safety and performance, advancing a more trustworthy and human-centered paradigm for shared autonomy.

2. Related Work

PROXIMAL STATE NUDGING builds on prior work in shared autonomy, cooperative AI, and human-centered approaches to trustworthy AI systems that support long-term human capability development.

Shared Autonomy and Cooperative AI Shared autonomy systems combine human and AI control to improve task execution while preserving human involvement. Early work introduced policy blending and assistance under uncertainty (Dragan & Srinivasa, 2013; Javdani et al., 2015), while later approaches leveraged deep reinforcement learning to relax assumptions about environment dynamics and user goals (Reddy et al., 2018). More recent work has explored learning-compatible assistance, where AI systems balance task performance with preserving opportunities for human learning (Bragg & Brunskill, 2020; Fitzsimons et al., 2019; Tian et al., 2023). Our work extends this direction by framing shared autonomy as a cooperative AI problem in which the system explicitly supports sustained human competence, not only immediate task success. Unlike prior approaches, PROXIMAL STATE NUDGING does not require access to the human learner's internal value function or an explicit model of human learning dynamics.

AI Systems for Human Skill Development A longstanding goal in human-centered AI is designing systems that actively help users learn and improve. Prior work has explored intelligent tutoring systems (Sleeman & Brown, 1982; Koedinger et al., 2015), skill assessment (Piech et al., 2015), and AI-driven coaching (Costa et al., 2025). In shared autonomy, recent approaches have drawn inspiration from educational psychology and the Zone of Proximal Development (ZPD) (Vygotsky, 1978) to guide users toward productive learning experiences (Agarwal & Deshpande, 2019; Srivastava et al., 2025). PROXIMAL STATE NUDGING similarly incorporates cognitive scaffolding principles, but does so through a general reinforcement learning framework that identifies and encourages transitions into high-learning-potential states while maintaining safety.

Skill Atrophy and AI Overreliance Growing evidence suggests that heavy reliance on AI assistance can lead to degradation of human skills and situational awareness. Skill atrophy has been documented in aviation (Casner et al., 2014), surgery (Sarofim, 2024), driving (de Winter et al., 2023b), and software engineering (Vaithilingam et al., 2022; Barke et al., 2023; Shen & Tamkin, 2026). Similar concerns have emerged around large language models in education and decision support (Kasneci et al., 2023). These findings highlight an important challenge for trustworthy AI: systems optimized solely for short-term performance may unintentionally reduce long-term human capability and resilience. Our work addresses this challenge directly by designing cooperative AI assistance that maintains strong performance while promoting continued human skill development.

3. Formalism

We consider a student with an unassisted policy $\pi_{\text{student},t} \in \Pi$ at each "learning timestep" t, mapping states to actions ($s \in S \to a \in A$). Our objective is to design a shared assistive policy $\pi_{\text{shared},t}$ that maps a state and the current student policy to an action ($s \in S, \pi_{\text{student},t} \in \Pi \to a \in A$), and satisfies two properties:

- **Useful assistance:** $\pi_{\text{shared},t}$ supports safe and effective task completion, measured by high reward.
- **Skill learning:** $\pi_{\text{student},t}$ improves over time, reflected by increasing unassisted reward over time.

Most shared autonomy methods prioritize the first property. For instance, assistive teleoperation often blends user inputs with an autonomous policy or overrides suboptimal actions (Wang et al., 2020; Reddy et al., 2018). However, determining how assistance should be structured to promote user learning remains challenging. Prior work models human adaptation with assumptions such as Wiener processes or linear-quadratic dynamics (Yu et al., 2023; Tian et al., 2023), limiting their ability to capture diverse learning behaviors.

In contrast, PROXIMAL STATE NUDGING adopts a state-centric view of skill learning, inspired by intelligent tutoring systems that operationalize Vygotsky's Zone of Proximal Development (ZPD) to adapt instruction difficulty (Vygotsky, 1978; Milani et al., 2020; Ropelato et al., 2018). We estimate a proximality function ϕ_{zpd} that measures how likely a state $s \in S$ is to support learning. States that are too difficult (little chance of reward) or too easy (already mastered) should have low proximality. The shared policy then nudges the student toward states with higher ϕ_{zpd} by selecting actions that trade off task optimality and state proximality. This joint optimization enables effective assistance while preserving unassisted learning.

3.1. Proximal State Estimation: Identifying Learnable States

A key cxomponent of PROXIMAL STATE NUDGING is estimating the state proximality function $\phi_{\text{zpd}}(s)$. ZPD is defined as the gap between what a learner can do independently and with assistance (Vygotsky, 1978). Accordingly, by collecting trajectories with ($\pi_{\text{shared},t}$) and without ($\pi_{\text{student},t}$) assistance, we approximate $\phi_{\text{zpd}}(s)$ as the difference in expected task reward:

$$\mathbb{E}_{\tau\sim\pi_{\text{shared},t},s\in\tau}[R_{\text{task}}(\tau)] - \mathbb{E}_{\tau\sim\pi_{\text{student},t},s\in\tau}[R_{\text{task}}(\tau)].$$

States from which task reward is similar with or without assistance — either because they are too easy or too hard — receive low $\phi_{\text{zpd}}(s)$. The ϕ_{zpd} estimator can be updated at each learning step and evolve with the student policy, and does not require an explicit student model. In practice, $\phi_{\text{zpd}}(s)$ is normalized, and function approximation is used for large or continuous state spaces (see Subsection 5.1).

3.2. Shared Control Planning

Given $\phi_{\text{zpd}}(s)$, we define $\pi_{\text{shared},t}$. We begin with a standard blending-based shared autonomy formulation (Wang et al., 2020; Steele & Gillespie, 2001):

$$\begin{aligned}&\pi_{\text{mix},t}(a \mid s) \triangleq \alpha\, \pi_{\text{agent}}(a \mid s) + (1-\alpha)\, \pi_{\text{student},t}(a \mid s),\\ &\pi_{\text{shared},t}(s) \triangleq \arg\max_{a\in\mathcal{A}} \pi_{\text{mix},t}(a \mid s).\end{aligned} \tag{1}$$

This policy selects the mode of a convex combination of the optimal agent and student policies, controlled by the assistance parameter α.

PROXIMAL STATE NUDGING modifies this by replacing the argmax with a beam search over a set of length-T sampled trajectories $\mathcal{B}_t(s) \sim \pi_{\text{mix},t}$. We then select the first action of the trajectory maximizing a score J, defined as a weighted sum of average proximality and task reward:

$$\begin{aligned}&\pi_{\text{mix},t}(a \mid s) \triangleq \alpha\, \pi_{\text{agent}}(a \mid s) + (1-\alpha)\, \pi_{\text{student},t}(a \mid s),\\ &J_t(\tau) \triangleq w_1 \left(\frac{1}{T}\sum_{i=0}^{T-1} \phi_{\text{zpd},t}(s_i)\right) + w_2\, R_{\text{task}}(\tau),\\ &\pi_{\text{shared},t}(s) \triangleq a_0(\tau^\star), \qquad \tau^\star \in \arg\max_{\tau\in\mathcal{B}_t(s)} J_t(\tau).\end{aligned} \tag{2}$$

Finally, we adapt α based on state proximality, reducing agent control in high-ϕ_{zpd} states where learning is most likely. This yields a shared policy that combines strong task assistance with targeted support for skill acquisition. Algorithm 1 summarizes our full approach; UpdateZPD and UpdatePolicy are task- and architecture-dependent.

Algorithm 1 Proximal State Nudging

```
Input: Task environment with state space $\mathcal{S}$, action space $\mathcal{A}$
    Optimal agent policy $\pi_{\text{agent}}$
    Assistance range $\alpha_{\min}, \alpha_{\max} \in [0,1]$
    Scoring weights $w_1, w_2 \in [0,1]$
    Planning horizon $T$
    Beam search size $B$
    Task reward function $R_{\text{task}}$
Initialize: Student policy $\pi_{\text{student},0}$, proximal state estimator $\phi_{\text{zpd}}$
for timestep $t$ do
    // Update proximal state estimator
    Collect assisted episodes: $\mathcal{D}_{\text{shared}} \sim \pi_{\text{shared},t}$
    Collect unassisted episodes: $\mathcal{D}_{\text{student}} \sim \pi_{\text{student},t}$
    $\phi_{\text{zpd},t} \leftarrow$ UpdateZPD($\mathcal{D}_{\text{shared}}, \mathcal{D}_{\text{student}}$)
    // Shared control execution
    for each encountered state $s$ do
        $\alpha_t(s) \leftarrow \max(\alpha_{\min}, \alpha_{\max} - \phi_{\text{zpd},t}(s))$
        $\pi_{\text{mix},t}(a \mid s) \leftarrow \alpha_t(s)\pi_{\text{agent}}(a \mid s) + (1-\alpha_t(s))\pi_{\text{student},t}(a \mid s)$
        $\mathcal{B}_t(s) \leftarrow$ BeamSearch($\pi_{\text{mix},t}, s, T, B$)
        for each trajectory $\tau = (s_0, a_0, \ldots, s_{T-1}) \in \mathcal{B}_t(s)$ do
            $J_t(\tau) \leftarrow w_1\left(\frac{1}{T}\sum_{i=0}^{T-1}\phi_{\text{zpd}}(s_i)\right) + w_2 R_{\text{task}}(\tau)$
        end for
        $\tau^\star \leftarrow \arg\max_{\tau\in\mathcal{B}_t(s)} J_t(\tau)$
        $\pi_{\text{shared},t}(s) \leftarrow a_0(\tau^\star)$
    end for
    // Update student policy
    $\pi_{\text{student},t+1} \leftarrow$ UpdatePolicy($\pi_{\text{student},t}$, experience)
end for
```

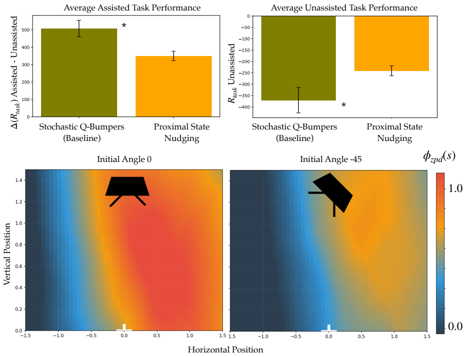


Figure 2: **Lunar Lander Task and Results** (Top) PROXIMAL STATE NUDGING results in higher unassisted task reward and a lower gap between unassisted and assisted task performance than Stochastic Q-Bumpers, averaged across 100 trials at $\alpha = 0.7$. Stat. sig. (*) determined with independent t-tests ($p < 0.05$). (Bottom) Heatmap showing estimated proximality scores for states in `Lunar Lander`, at different initial angles of the lander.

4. Empirical Results: Simulated Learners

We next evaluate PROXIMAL STATE NUDGING in with simulated learner agents. Specifically, we focus on a 6-state `DiscreteMDP` (discrete action) and the `Lunar Lander` (continuous action) task studied in Bragg & Brunskill (2020), which proposed Stochastic Q-Bumpers, a method for learning-compatible decision support. We follow Bragg & Brunskill (2020) by analyzing the pareto frontier between two metrics: fully assisted and unassisted task performance. We simulate the student agents in both `DiscreteMDP` and `Lunar Lander` as Q-Learning agents, using DQN for `Lunar Lander` (Mnih et al., 2015).

4.1. Simple Discrete MDP

Task Set-Up We consider a 6-state Markov Decision Process with a start state s_0, goal state s_g, and intermediate states $s_1, \dots, s_5$. The action space consists of 5 discrete actions, with transitions defined in Figure 3. To increase difficulty, transitions from all states except s_0 and s_g remain in the same state with probability $p = 0.2$. The agent always starts at s_0 and receives reward $r = 1$ at s_g, and a small step penalty $r = -0.1$ otherwise. With horizon $H = 5$, the MDP is structured so that some states (e.g., s_2) are easier than others (e.g., s_5) for reaching s_g.

Methods We study a tabular Q-learning student under either **Fully-Assisted Shared Autonomy** ($\alpha\,\pi_{\text{agent}} + (1-\alpha)\,\pi_{\text{student}}$) or our **PROXIMAL STATE NUDGING** approach, where π_{agent} corresponds to the optimal policy. For PROXIMAL STATE NUDGING, we follow Algorithm 1 with $w_1 = 1$, $w_2 = 1$, and $T = 2$, appropriate for the short horizon. We compare the ability of both methods to balance *assisted* and *unassisted* ($\alpha = 0$) performance, where improved unassisted reward indicates student learning.

Results Figure 3 reports assisted and unassisted rewards over 10 time steps for different α. Both methods achieve high assisted performance. However, under **Fully-Assisted Shared Autonomy**, unassisted performance quickly plateaus at a low reward, indicating that assistance impedes independent learning. In contrast, PROXIMAL STATE NUDGING yields steadily improving unassisted performance, showing that nudging based on ϕ_{zpd} preserves the student's ability to learn. The ϕ_{zpd} values at $t = 50$ and $\alpha = 0.7$ (Figure 3, left) indicate low scores for s_2, where assistance is unnecessary, and higher scores for s_1 and s_3, which are prioritized for nudging. Overall, across α, PROXIMAL STATE NUDGING improves unassisted performance while maintaining strong assisted performance, demonstrating benefits beyond simply reducing assistance.

4.2. Lunar Lander

Task Set-Up `LunarLander` (Brockman et al., 2016) is a classic continuous control benchmark in which an agent must safely land a spacecraft in a 2D environment. The environment has a continuous 8-dimensional state space encoding position, velocity, orientation, angular velocity, and ground contact indicators, and a discrete action space of four engine commands. The reward function is shaped to encourage proximity to the landing pad, low velocities, upright orientation, and successful leg contact, while penalizing fuel usage and crashes. Further experimental details, including hyperparameters, are in the Appendix.

Methods We compare our **PROXIMAL STATE NUDGING** method, where π_{agent} is set to an optimal agent we trained using DQN (Mnih et al., 2015), with the Stochasic Q-Bumpers agorithm proposed by Bragg & Brunskill (2020). Stochastic Q-Bumpers stochastically overrides a user's input with the optimal action whenever the Q-value of the user's selected action strongly differs from the optimal Q value. The likelihood of overriding is proportional to the degree the user differed from the optimal policy, and thus this approach is able to preserve some learning of the task while still preventing the overall episode reward from decreasing. Bragg & Brunskill (2020) showed that Stochastic Q-Bumpers outperforms other forms of shared autonomy, such as (Reddy et al., 2018), which is why we focus specifically on comparing with this method for `Lunar Lander`.

Results Figure 2 shows that overall unassisted task performance is significantly higher with **PROXIMAL STATE NUDGING**, while maintaining a significantly lower gap between unassisted and assisted task performance than Stochastic Q-Bumpers. In practice, we found our implementation of Stochastic Q-Bumpers to be unstable to train, and hypothesize that this is because the method lacks an explicit mechanism for identifying and targeting states within the learner's zone of proximal development. By only preventing large deviations from optimality, it may keep learners in overly safe regions or allow them into states too difficult for their current skill level. Our proximal score estimator ϕ_{zpd} explicitly identifies states where learning is maximized, enabling more targeted guidance. A closer look at the outputs of a trained $\phi_{zpd}(s)$ model across different positions of the Lunar Lander environment (fixing the velocity to the average across trajectories drawn from π_{agent}), shows that the model assigns highest proximality scores to states where the lander has moderate angular displacement. This indicates that PROXIMAL STATE NUDGING preferentially guides learners toward states requiring active stabilization and orientation control, rather than keeping them in trivially stable configurations.

Our simulation results so far assumed the student agent

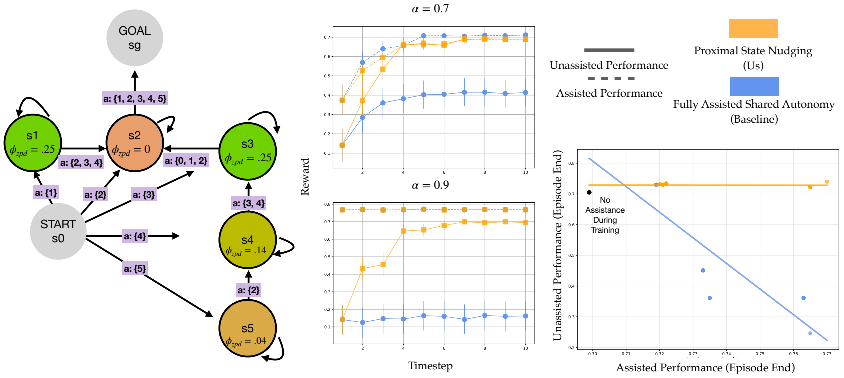


Figure 3: **Discrete MDP Task and Results (Best Viewed in Color).** (Left) MDP with ϕ_{zpd} at $t = 50$ ($\alpha = 0.7$), illustrating that PROXIMAL STATE NUDGING guides the agent toward s_1 and s_3 to support unassisted learning. (Middle) Learning curves comparing both methods for two assistance levels (α), averaged over 50 trials. (Right) Unassisted versus assisted reward across α, showing that PROXIMAL STATE NUDGING achieves a superior Pareto frontier.

employed Q-learning updates (either tabular or with function approximation). What happens when we evaluate with real human students, who may not follow a straightforward learning strategy?

5. Empirical Results: Human Subjects

We next evaluate how PROXIMAL STATE NUDGING supports real human users learning two dynamical control tasks: Parallel Parking and High Performance Racing. We built environments for both tasks using the CARLA simulator (Dosovitskiy et al., 2017), often used for autonomous driving research (Li et al., 2023; Srivastava et al., 2025; Costa et al., 2025; Mihaylova et al., 2025). We ask two research questions:

1. **RQ1 (Learning)** Does PROXIMAL STATE NUDGING improve learning (change in unassisted task performance)? We hypothesize that PSN will outperform *fully-assisted shared autonomy*, and comparable to a *self-practice* baseline.
2. **RQ2 (Safety)** Does PROXIMAL STATE NUDGING maintain strong overall task performance during learning? We hypothesize that will outperform *self-practice*, while remaining comparable to *fully-assisted shared autonomy*.

The unifying hypothesis is that PSN strikes a balance between learning gains (unassisted) and safe task performance during learning (assisted) when compared to baselines. Since we anticipate that learning is a relatively slower temporal process for human studenrs than for simulated students, we build our ϕ_{zpd} estimator using offline data, assuming we can treat state proximality as fixed for the 1 hour duration of our experiments. We recruit a total of $n = 60$ participants, mostly university students, across two institutions. Our study was approved by our Institutional Review Board, and we provide more details about study instructions and post-survey questions in the Appendix.

5.1. High Performance Racing

In the high-performance racing (HPR) task, users are instructed to drive around a fixed race track with the goal of minimizing their overall lap time. This task is particularly well-suited for studying human skill acquisition because it has clear success metrics for both learning (e.g. improved lap time) and safety (e.g. total number of collisions).

5.1.1. TASK OVERVIEW

We follow the HPR task set-up studied in several prior works (DeCastro et al., 2024; Lidard et al., 2024; Srivastava et al., 2025; Sumner et al., 2025). Specifically, we we simulate the 2-mile track in Thunderhill Raceway Park, California in the CARLA simulator, using a vehicle configured for high performance racing. Each state in the HPR task consists of the vehicle's position (x, y, z), angle (yaw, $pitch$, $roll$), and speed (vx, vy, vz). The continuous action space consists of steering angle, throttle, and brake control. Each episode (a single lap around the raceway) is evaluated with three metrics: lap time, Dynamic Time Warp distance from an expert demonstration, and smoothness (measured as jerk action input).

Methods We compare three methods:

- **Fully-Assisted Shared Autonomy (Baseline)**: We blend $\pi_{student}$, the user's input controls, and π_{agent}, a policy based on an expert human demonstration, with $\alpha = 0.8$ following Equation 1.

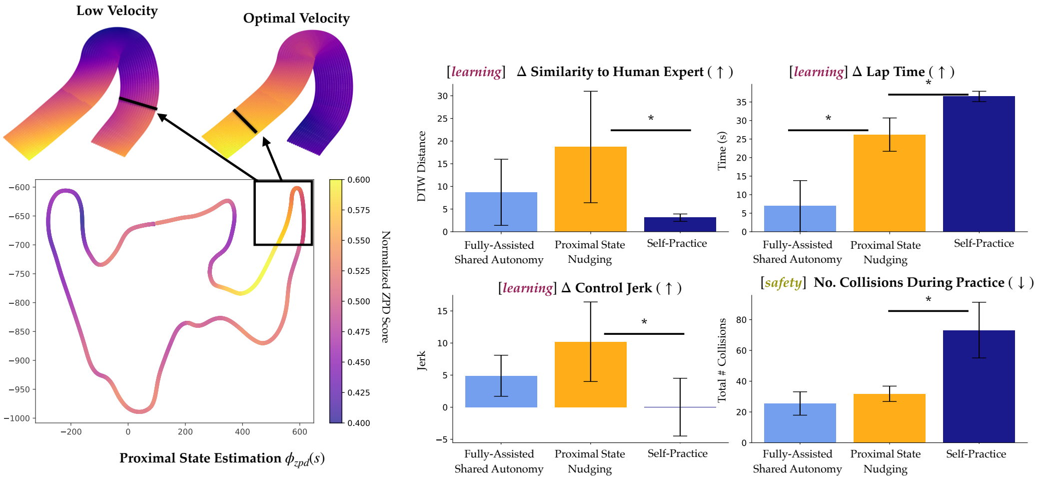


Figure 4: **High Performance Racing Task and Results** (Left) Heatmap of the learned state proximality estimator ϕ_{zpd} shows that assistance is predicted to most strongly support student learning at states around high-speed tight turns, and that the proximality estimates are sensitive to both location and velocity. (Right) PROXIMAL STATE NUDGING successfully balances the trade-off between learning gains (e.g. improved similarity to expert, lap time, and control jerk) and safety during practice, in comparison to both fully-assisted shared autonomy and unassisted self-practice baselines. Bars report mean standard error, statistical significance determined with Welch's t-test.

- **Proximal State Nudging ()**: We sample, at fixed intervals along the track, the nearest point for each of 5 distinct human expert trajectories, which cover different near-optimal task strategies. We then set the point in each trajectory that crosses the start of the next interval as the goal for CARLA's Global Route Planner (GRP), which returns a set of waypoint trajectories $\mathcal{B}_t(s)$. Following Equation 2, we set R_{task} to the normalized inverse trajectory length (encouraging high speed), and $w_1 = 0.5$ and $w_2 = 0.5$.

- **Unassisted Self-Practice (Baseline)**: The student repeatedly practices high performance racing around the track with full control of the vehicle.

5.1.2. TECHNICAL IMPLEMENTATION

Assistive Control Design As described above, the assistive control for both the Fully-Assisted Shared Autonomy baseline and use demonstrations from a human expert, which were collected by (Srivastava et al., 2025) from a member of an HPR club experienced with the Thunderhill raceway. Given a point in an expert trajectory, the built-in Path Planner in the CARLA simulator generates waypoints to input into a PID controller with the driver's state, and the control outputs are used as the actions for π_{agent}.

Proximal State Estimation To avoid introducing latency for our human users, we used a Linear Regression model for ϕ_{zpd}. This is suitable as the state space is low-dimensional. To train ϕ_{zpd}, we collect an offline dataset of 50 demonstrations of novice drivers both with and without assistance (at $\alpha = 0.8$). We then train *two* separate linear regression models (ρ_{sa} and ρ_{ua}) to predict the segment reward, corresponding each to the shared-autonomy and unassisted conditions, and then define the overall estimator using an affine normalization of the output difference ($\rho_{\text{sa}}(s) - \rho_{\text{ua}}(s)$). This two-model approach is necessary due to the continuous state space resulting in trajectories that rarely overlap exactly.

Figure 4 (left) shows a heatmap of the outputs of ϕ_{zpd} across different states across the Thunderhill Raceway track, showing that the states with the highest estimated learning potential for students are concentrated at the tight turn towards the end of the track. Furthermore, $\phi_{zpd}(s)$ is very sensitive to the velocity values in s, suggesting that the student's speed affects its proximality. Interestingly, $\phi_{zpd}(s)$ changes laterally across the track, and it appears PROXIMAL STATE NUDGING would nudge a slower student to take the tight turn widely, while a faster student would be encouraged to focus on the execution after the turn's apex.

5.1.3. HUMAN SUBJECT STUDY RESULTS

We compare all three methods with the same sequence of trials:

1. 2 **baseline trials** with unassisted user control.
2. 2 **practice trials** with either (i) Fully Assisted Shared Autonomy, (ii) Unassisted Self-Practice or (iii) PROXIMAL STATE NUDGING.
3. 2 **evaluation trials** with unassisted user control.

In each trial, subjects were instructed to complete one lap as quickly as possible (with a 3 minute limit). We recruited 30 novice drivers, evenly split between methods, with 90% holding a driver's license. We compute the three learning metrics (lap time, Dynamic Time Wrap distance to a human expert, and control input jerk) as the difference between average baseline and evaluation performance, where larger values indicate greater gains. For safety, we measure the total number of collisions during practice trials, with lower values indicating safer learning.

Our results, reported in Figure 4, show that PROXIMAL STATE NUDGING outperformed Fully Assisted Shared Autonomy in improvement from baseline to evaluation trials across all three learning metrics, with a statistically significant difference in lap time, supporting our hypothesis for **RQ1**. was also significantly stronger than Unassisted Self-Practice for both similarity to expert and jerk metrics, though self-practice led to a significantly stronger improvement in lap time. Finally, users assigned either or Fully Assisted Shared Autonomy conditions had low number of collisions during practice trials, but significantly lower than Unassisted Self-Practice, validating our hypothesis for **RQ2**.

One possible reason why self-practice leads to a significant improvement in lap time, but not other metrics, is that students are so fixated on improving the lap time that they ignore other aspects of good driving like smooth control. In fact, participants assigned during practice noting learning specific turning skills (e.g., *"I learned there are certain chokepoints where you need to reduce speed to avoid spinning out"*). Furthermore, students doing self-practice had their learning gains come at a significant cost to safety, as the number of collisions during practice was high. Therefore, these results support our hypotheses that PROXIMAL STATE NUDGING successfully improves student's learning (**RQ1**) while still providing an appropriate degree of assistance to ensure safety (**RQ2**).

5.2. Parallel Parking

Parallel parking is another challenging driving task requiring precise spatial reasoning, planning, and motor control. It is commonly used in autonomous driving research due to its difficulty for human drivers and clear success criteria (Zhang et al., 2020; Srivastava et al., 2022; Chai et al., 2023). The task consists of an approach phase followed by a parking phase, and we assume learning difficulties primarily arise from variation in the endpoint of the approach phase (i.e., the start of the parking phase).

5.2.1. TASK OVERVIEW

We again use the CARLA simulator, this time implementing our parallel parking in the provided CARLA Town15 multilane suburban environment. Each state includes vehicle position (x, y, z), orientation, and velocity, while continuous actions consist of steering, throttle, and brake. We designated a 13 m gap between two stationary vehicles as the target spot, and each episode (a single parking attempt) is evaluated with two metrics: parking time and number of collisions.

Methods We compare three methods:

- **Fully-Assisted Shared Autonomy (Baseline):** We belnd $\pi_{student}$, the user's input controls, and π_{agent}, a pure-pursuit controller that takes a goal state as input, with $= 0.8$ following Equation 1.
- **Proximal State Nudging ()** : We sample different states for the start of the parking phase, and we set the state with the highest ϕ_{zpd} estimate as the goal for a pure pursuit controller that guides the student. Following Equation 2, we set R_{task} to a fixed constant, and $w_1 = .5$ and $w_2 = .5$.
- **Unassisted Self-Practice (Baseline)**: The student repeatedly practices parallel parking with full control of the car.

5.2.2. TECHNICAL IMPLEMENTATION

Assistive Control Design For our assistive control, the optimal agent π_{agent} uses a non-linear Model Predictive Control (NMPC) policy: a proportional heading-and-speed controller for the approach, followed by a finite-horizon nonlinear program (solved with IPOPT) for parking (Wächter & Biegler, 2006). The NMPC controls are blended with human steering inputs using a fixed mixing ratio $\alpha = 0.8$, and the agent is queried every 8 frames for tractability.

Proximal State Estimation We discretize the high-level state space near the parking spot and estimate ZPD scores for each stat. We collect 272 trajectories from 17 novice drivers, each attempting 4 trials from 8 spots (2 assisted, 2 unassisted, randomized). FollowingSection 3, $\phi_{zpd}(s)$ is computed as the average difference of $\rho_{sa}(s)$ and $\rho_{ua}(s)$, which denote average parking time for assisted and unassisted trials . Figure 5 (left) shows the state directly in front of the target spot has the highest estimated learning potential.

5.2.3. HUMAN SUBJECT STUDY RESULTS

We compare all three moethods with the same trial order:

1. 2 **baseline trials** with unassisted user control.
2. 4 **practice trials** with either (i) Fully Assisted Shared Autonomy, (ii) Unassisted Self-Practice, or (iii) PROXIMAL STATE NUDGING.

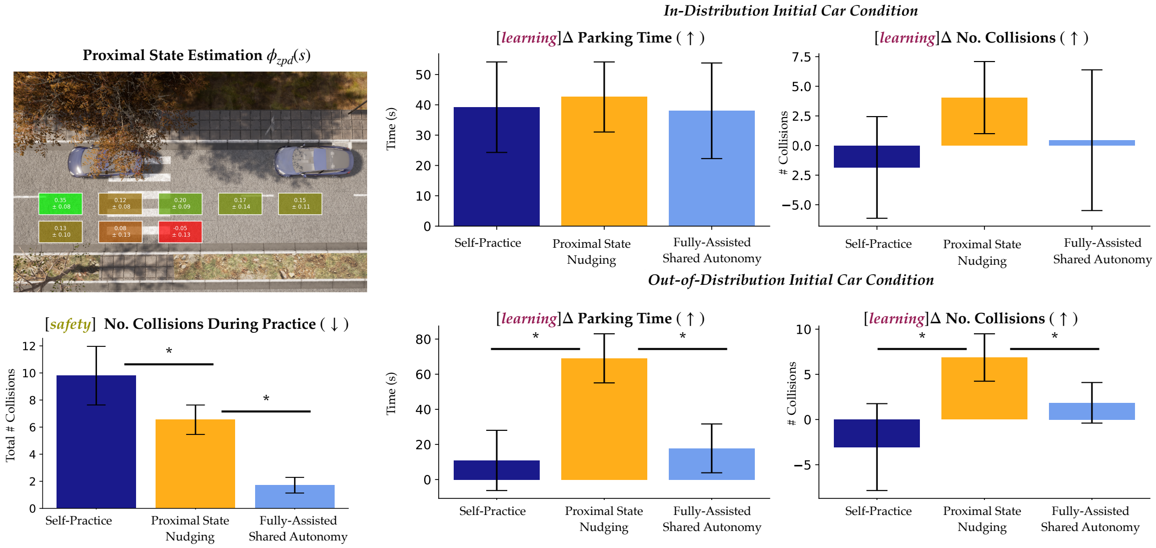


Figure 5: **Parallel Parking Task and Results** (Top Left) Estimated proximal learning potential ϕ_{zpd} over discretized approach states near the target parking spot; green indicates higher learning potential. (Bottom Left and Right) PROXIMAL STATE NUDGING succesfully balances the trade-off betwen learning gains (e.g. improved parking time and collisions) and safety (total number of collisions during practice). in comparison to both fully-assisted shared autonomy and unassisted self-practice baselines. Bars show mean standard error, statistical significance via two-sample Welch's t-tests.

3. 2 **evaluation trials** with unassisted user control (in distribution starting angle)
4. 2 **evaluation trials** with unassisted control (out of distribution starting angle not seen during baseline trials)

In each trial, participants were instructed to to park as quickly as possible while avoiding collisions. We recruited 30 novice drivers, evenly split between conditions, all holding a driver's license. We compute two learning metrics (parking time and of collisions) as the difference between average baseline and evaluation performance, so larger values indicate stronger learning gains. FOr safety, we measure the total number of collisions during practice trials, with lower values indicate safer learning.

Our results, reported in Figure 5, show that **Proximal State Nudging** yields greater improvement than both unassisted self-practice and fully-assisted shared autonomy in both improving parking time and reducing collisions, support our hypothesis for RQ1. However, statistically strong effects were primarily observed for the randomized-angle out-of-distribution starting angle, which suggests that PROXIMAL STATE NUDGING's effect is more salient in conditions that are unfamiliar to a student. Furthermore, users assigned PROXIMAL STATE NUDGING has significantly fewer collisions during practice trials than Unassisted Self-Practice which validates our hypothesis for **RQ2**, although Fully Assisted Shared Autonomy did result in an even more significantly safer practice.

Participants who practices without assistance mainly reported self-learning steering wheel sensitivity, while participants more often described learning new parking strategies (e.g., *"I came to better understand where I should turn my wheel."*). Overall, the results for our Parallel Parking task support our prior experimental results in affirming the goal of PROXIMAL STATE NUDGING to provide assistance that can ensure safety without compromising users' learning, thereby mitigating skill atrophy.

6. Limitations and Future Work

Our results suggest that PROXIMAL STATE NUDGING can balance effective AI assistance with sustained human skill development across both simulated and human-subject shared autonomy tasks. While our studies demonstrate promising improvements in learning-compatible assistance, future work should validate these findings with larger and more diverse participant populations, investigate long-term retention effects, and evaluate transfer to deployment.

More broadly, the challenge of preventing skill atrophy extends beyond control domains to many forms of AI assistance, including education, coding, and decision support. Future work could explore how cooperative AI systems can adaptively regulate assistance to preserve human agency, competence, and resilience across these settings.

References

Agarwal, P. and Deshpande, A. A framework for adaptation of training task, assistance and feedback for optimizing motor (re)-learning with a robotic exoskeleton. *IEEE Robotics and Automation Letters*, 4:808–815, 2019. doi: 10.1109/LRA.2019.2891431.

Barke, S., James, M. B., and Polikarpova, N. Grounded copilot: How programmers interact with code-generating models. *Proceedings of the ACM on Programming Languages*, 7(OOPSLA1):85–111, 2023.

Bragg, J. and Brunskill, E. Fake it till you make it: Learning-compatible performance support. In Adams, R. P. and Gogate, V. (eds.), *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pp. 915–924. PMLR, 22–25 Jul 2020. URL https://proceedings.mlr.press/v115/bragg20a.html.

Brockman, G., Cheung, V., Pettersson, L., et al. OpenAI gym. *arXiv preprint arXiv:1606.01540*, 2016.

Casner, S. M., Hutchins, E. L., and Norman, D. The impact of automation on pilot cognition and performance. *Human Factors*, 56(1):1–17, 2014.

Casner, S. M., Hutchins, E. L., and Norman, D. The challenges of partially automated driving. *Communications of the ACM*, 59(5):70–77, 2016.

Chai, R., Liu, D., Liu, T., Tsourdos, A., Xia, Y., and Chai, S. Deep learning-based trajectory planning and control for autonomous ground vehicle parking maneuver. *IEEE Transactions on Automation Science and Engineering*, 20 (3):1633–1647, 2023. doi: 10.1109/TASE.2022.3183610.

Costa, J., Morgan, A., Yasuda, H., Sumner, E. S., Gopinath, D., Chau, S., Nguyen, H., Best, A., Rosman, G., and Chen, T. L. From dashboards to dialogue: Evaluating a conversational ai coach for performance driving skill development. In *Proceedings of the 17th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '25, pp. 9–18, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400720130. doi: 10.1145/3744333.3747811. URL https://doi.org/10.1145/3744333.3747811.

de Winter, J. C., Petermeijer, S. M., and Abbink, D. A. Shared control versus traded control in driving: a debate around automation pitfalls. *Ergonomics*, 66(10):1494–1520, 2023a.

de Winter, J. C. F., Petermeijer, S. M., and Abbink, D. A. Shared control versus traded control in driving: a debate around automation pitfalls. *Ergonomics*, 66(10):1494–1520, October 2023b.

DeCastro, J., Silva, A., Gopinath, D., Sumner, E., Balch, T. M., Dees, L., and Rosman, G. Dreaming to assist: Learning to align with human objectives for shared control in high-speed racing, 2024. URL https://arxiv.org/abs/2410.10062.

Dell'Aquila, E., Rossi, F., and Ronchetti, M. Chatgpt in education: transforming digital learning and assessment. *Applied Sciences*, 13(18):10053, 2023.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. Carla: An open urban driving simulator. In *Proceedings of the 1st Annual Conference on Robot Learning (CoRL)*, pp. 1–16, 2017.

Dragan, A. D. and Srinivasa, S. S. A policy-blending formalism for shared control. *International Journal of Robotics Research (IJRR)*, 32:790–805, 2013.

Endsley, M. R. Autonomous driving systems: A preliminary naturalistic study of the tesla model s. *Journal of Cognitive Engineering and Decision Making*, 11(3):225–238, 2017.

Fitzsimons, K., Kalinowska, A., Dewald, J., and Murphey, T. Task-based hybrid shared control for training through forceful interaction. *The International Journal of Robotics Research*, 39:1138 – 1154, 2019. doi: 10.1177/0278364920933654.

Javdani, S., Srinivasa, S., and Bagnell, J. A. Shared autonomy via hindsight optimization. In *Robotics: Science and Systems*, 2015.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., et al. Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and individual differences*, 103: 102274, 2023.

Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., and Bier, N. L. Learning is not a spectator sport: Doing is better than watching for learning from a mooc. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, L@S '15, pp. 111–120, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450334112. doi: 10.1145/2724660.2724681. URL https://doi.org/10.1145/2724660.2724681.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

Li, Z., Sun, Y., Tomizuka, M., and Zhan, W. Learning from active human involvement through proxy value propagation. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, 2023.

Lidard, J., Hu, H., Hancock, A., Zhang, Z., Contreras, A. G., Modi, V., DeCastro, J., Gopinath, D., Rosman, G., Leonard, N. E., et al. Blending data-driven priors in dynamic games. *arXiv preprint arXiv:2402.14174*, 2024.

Macnamara, B. N., Berber, I., Çavuşoğlu, M. C., Krupinski, E. A., Nallapareddy, N., Nelson, N. E., Smith, P. J., Wilson-Delfosse, A. L., and Ray, S. Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? *Cognitive Research: Principles and Implications*, 9(1):46, 2024.

Mihaylova, T., Reitmann, S., Topp, E. A., and Kyrki, V. Injecting conflict situations in autonomous driving simulation using carla. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 1052–1056, 2025. doi: 10.1109/HRI61500.2025.10974213.

Milani, S., Fan, Z., Gulati, S., Nguyen, T., Fang, F., and Yadav, A. Intelligent tutoring strategies for students with autism spectrum disorder: A reinforcement learning approach. In *The 2020 CMU Symposium on Artificial Intelligence and Social Good*, 2020.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540): 529–533, 2015.

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., and Sohl-Dickstein, J. Deep knowledge tracing. *Advances in neural information processing systems*, 28, 2015.

Reddy, S., Dragan, A. D., and Levine, S. Shared autonomy via deep reinforcement learning. In *Arxiv 1802.01744*, 2018. URL https://arxiv.org/abs/1802.01744.

Ropelato, S., Zünd, F., Magnenat, S., Menozzi, M., and van Dinther, Y. Adaptive tutoring on a virtual reality driving simulator. *ETH ZurichInternational SERIES on Information Systems and Management in Creative EMedia*, 2017 (2):12–17, 2018.

Sarofim, M. Devil's advocate: exploring the potential negative impacts of artificial intelligence on the field of surgery. *Journal of Medical Artificial Intelligence*, 7, 2024.

Shen, J. H. and Tamkin, A. How ai impacts skill formation. *arXiv preprint arXiv:2601.20245*, 2026.

Sleeman, D. and Brown, J. S. *Intelligent tutoring systems*. London: Academic Press, 1982.

Srivastava, M., Biyik, E., Mirchandani, S., Goodman, N., and Sadigh, D. Assistive teaching of motor control tasks to humans. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A. (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 28517–28529. Curran Associates, Inc., 2022.

Srivastava, M., Iranmanesh, R., Cui, Y., Gopinath, D., Sumner, E. S., Silva, A., Dees, L., Rosman, G., and Sadigh, D. Shared autonomy for proximal teaching. In *Proceedings of the 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2025)*, pp. 232–241, 2025.

Steele, M. and Gillespie, R. B. Shared control between human and machine: Using a haptic steering wheel to aid in land vehicle guidance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45(23):1671–1675, 2001. doi: 10.1177/154193120104502323. URL https://doi.org/10.1177/154193120104502323.

Sumner, E., Gopinath, D. E., Dees, L., Gomez, P. R., Cui, X., Silva, A., Costa, J., Morgan, A., Schrum, M., Chen, T. L., Balachandran, A., and Rosman, G. Simcoachcorpus: A naturalistic dataset with language and trajectories for embodied teaching, 2025. URL https://arxiv.org/abs/2509.14548.

Tian, R., Bajcsy, A., Tomizuka, M., and Dragan, A. D. Towards modeling and influencing the dynamics of human learning. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 12–21. Association for Computing Machinery, 2023. ISBN 978-1-4503-9973-7. doi: 10.1145/3568162.3578629.

Vaithilingam, P., Zhang, T., and Glassman, E. L. Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models. In *Chi conference on human factors in computing systems extended abstracts*, pp. 1–7, 2022.

Vygotsky, L. S. *Mind in society: The development of higher psychological processes*, volume 86. Harvard university press, 1978.

Wächter, A. and Biegler, L. T. On the implementation of a primal-dual interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006. doi: 10.1007/s10107-004-0559-y.

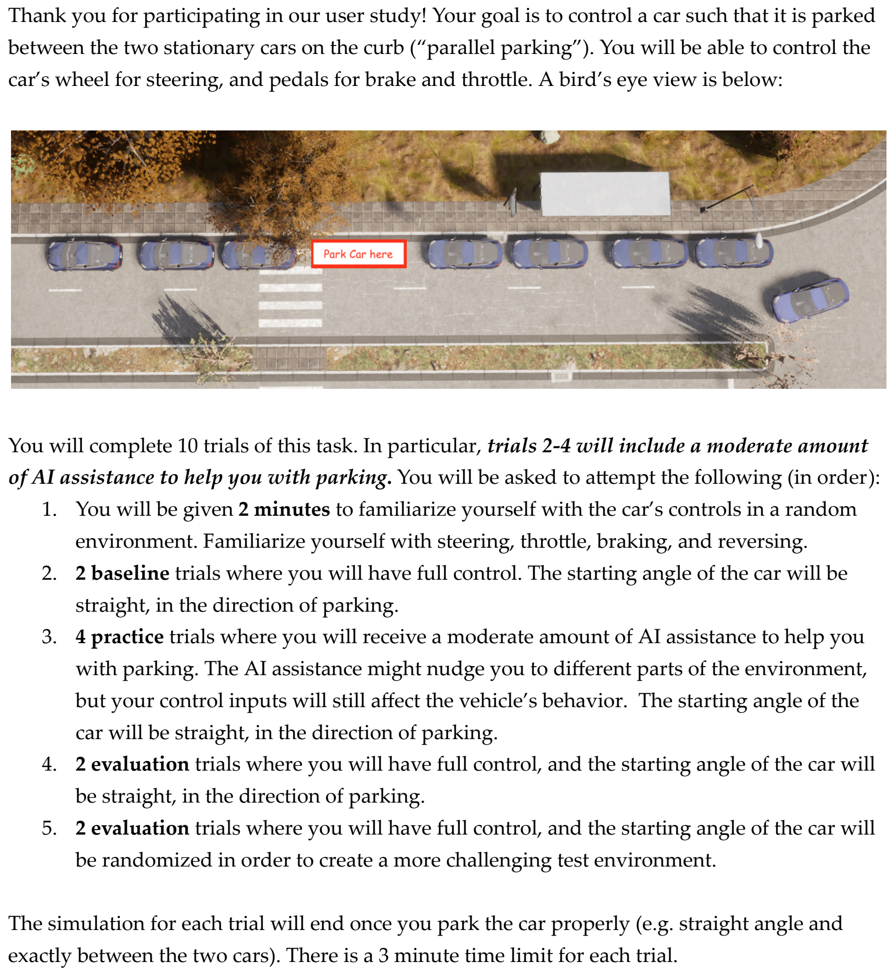
Wang, W., Na, X., Cao, D., Gong, J., Xi, J., Xing, Y., and Wang", F.-Y. Decision-making in driver-automation shared control: A review and perspectives. *IEEE/CAA Journal of Automatica Sinica*, 7(JAS-2020-0177):1289, 2020. ISSN 2329-9266. doi: 10.1109/JAS.2020.1003294.

Yu, C., Xu, Y., Li, L., and Hsu, D. Coach: Cooperative robot teaching. In Liu, K., Kulic, D., and Ichnowski, J. (eds.), *Proceedings of The 6th Conference on Robot Learning*, volume 205 of *Proceedings of Machine Learning Research*, pp. 1092–1103. PMLR, 14–18 Dec 2023. URL `https://proceedings.mlr.press/v205/yu23b.html`.

Zhang, J., Chen, H., Song, S., and Hu, F. Reinforcement learning-based motion planning for automatic parking system. *IEEE Access*, 8:154485–154501, 2020. doi: 10.1109/ACCESS.2020.3017770.

A. Appendix

A.1. User Study Instructions



Thank you for participating in our user study! Your goal is to control a car such that it is parked between the two stationary cars on the curb (“parallel parking”). You will be able to control the car’s wheel for steering, and pedals for brake and throttle. A bird’s eye view is below:

You will complete 10 trials of this task. In particular, ***trials 2-4 will include a moderate amount of AI assistance to help you with parking.*** You will be asked to attempt the following (in order):

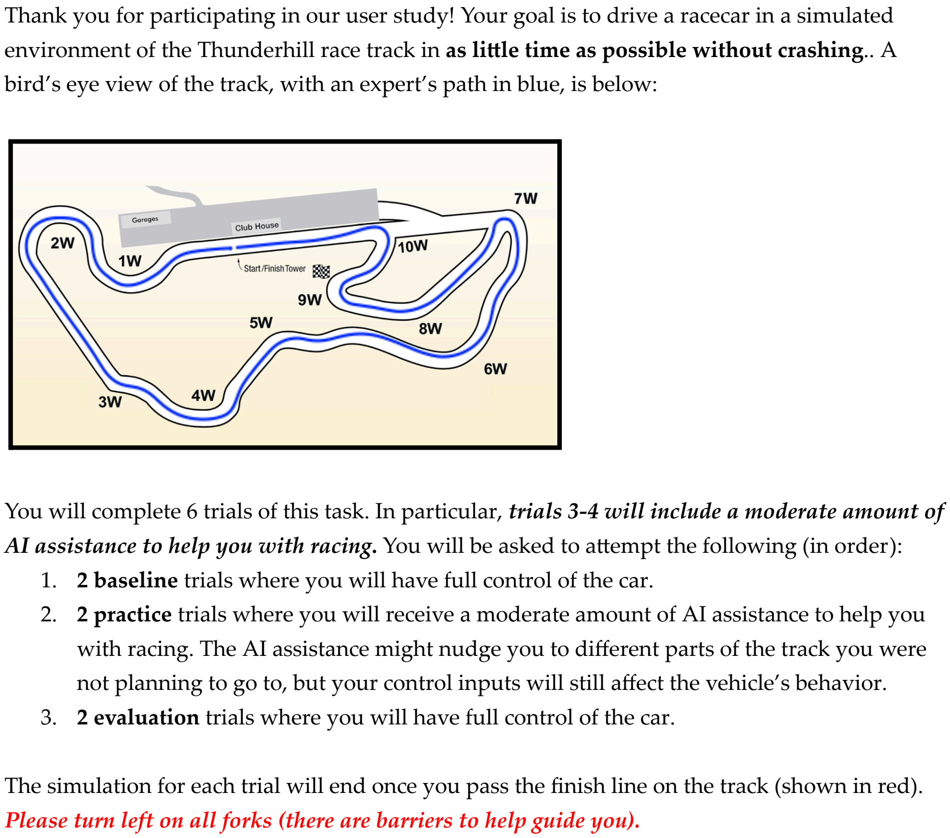
1. You will be given **2 minutes** to familiarize yourself with the car’s controls in a random environment. Familiarize yourself with steering, throttle, braking, and reversing.
2. **2 baseline** trials where you will have full control. The starting angle of the car will be straight, in the direction of parking.
3. **4 practice** trials where you will receive a moderate amount of AI assistance to help you with parking. The AI assistance might nudge you to different parts of the environment, but your control inputs will still affect the vehicle’s behavior. The starting angle of the car will be straight, in the direction of parking.
4. **2 evaluation** trials where you will have full control, and the starting angle of the car will be straight, in the direction of parking.
5. **2 evaluation** trials where you will have full control, and the starting angle of the car will be randomized in order to create a more challenging test environment.

The simulation for each trial will end once you park the car properly (e.g. straight angle and exactly between the two cars). There is a 3 minute time limit for each trial.

Figure 6: User study instructions for parallel parking task.

A.2. `LunarLander` Training Details

π_{agent} is a pre-trained DQN-based policy trained for 100 episodes, with learning rate .001, and an MLP-archicture consisting of 2 linear layers and hidden dimension 64. ϕ_{zpd} is also an MLP with 3 linear layers of dimensions 32, 32, and 16 respectively. All evaluation results in the paper are averaged across 100 trials.



Thank you for participating in our user study! Your goal is to drive a racecar in a simulated environment of the Thunderhill race track in **as little time as possible without crashing**.. A bird's eye view of the track, with an expert's path in blue, is below:

You will complete 6 trials of this task. In particular, ***trials 3-4 will include a moderate amount of AI assistance to help you with racing.*** You will be asked to attempt the following (in order):

1. **2 baseline** trials where you will have full control of the car.
2. **2 practice** trials where you will receive a moderate amount of AI assistance to help you with racing. The AI assistance might nudge you to different parts of the track you were not planning to go to, but your control inputs will still affect the vehicle's behavior.
3. **2 evaluation** trials where you will have full control of the car.

The simulation for each trial will end once you pass the finish line on the track (shown in red). *Please turn left on all forks (there are barriers to help guide you).*

Figure 7: User study instructions for HPR task.