Supplementary

This supplementary material presents additional simulation results, and provides additional details on the human subjects study in simulation as well as the in-person user study with a Kinova arm.

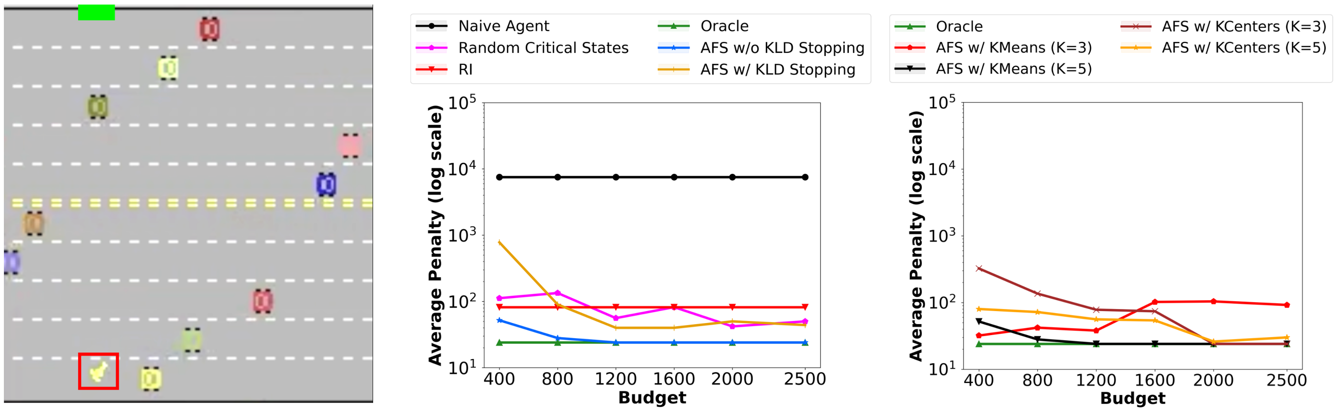
1 ADDITIONAL EXPERIMENTS IN SIMULATION

1.1 Learning with Implicit and Explicit Feedback Formats

To demonstrate AFS's flexibility across feedback modalities, we extend it to handle both explicit (e.g., DAM) and implicit (e.g., gaze) feedback formats.

1.1.1 Implicit Feedback Format

We consider a human's gaze on the screen as the implicit feedback format. Here, the robot requests to collect gaze data of the user and compares its action outcomes with the gaze positions of the user (Saran et al., 2021). Actions with outcomes aligning with the average gaze direction are labeled as acceptable (l_a), and unacceptable (l_h) otherwise.



(a) ATARI Freeway: Avoidable NSE (b) Average penalty incurred (c) Average penalty incurred with KMeans and KCenters clustering.

Figure S1: An instance of the Freeway domain, and the average penalty incurred.

1.1.2 Domain

We evaluate AFS in the Atari Freeway environment, where the robot (a chicken) navigates ten cars moving at varying speeds to reach the destination quickly while avoiding being hit(Figure S1(a)). Being hit by a car moves the robot back to its previous position, and is a severe NSE. A game state is defined by coordinates (x_1, y_1) and (x_2, y_2), i.e., the top left and bottom right corners of the robot and cars, extracted from the Atari-HEAD dataset (Zhang et al., 2020). Similar to (Saran et al., 2021), only car coordinates within a specific range of the robot are considered. The robot can move up, down or stay in place, with unit cost and deterministic transitions.

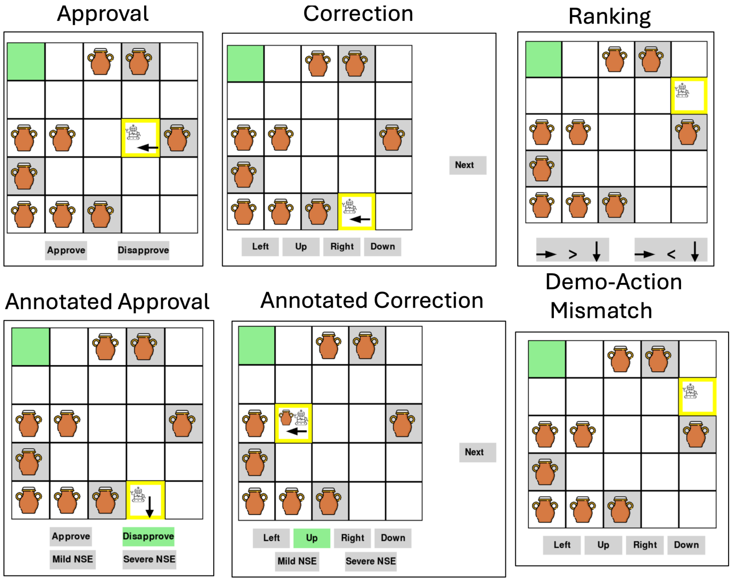


Figure S2: Interface for the human subjects study in simulation. Participants provide feedback via button clicks, with available options varying by format, as shown.

1.1.3 Results and Discussion

Effect of Learning using AFS. Figure S1(b) shows the average NSE penalties when operating based on an NSE model learned using different querying approaches. Clusters for critical state selection were generated using KMeans clustering algorithm with $K = 5$ in the Atari Freeway domain. Table S1 shows the average cost for task completion. While the Naive Agent has a lower cost for task completion, it incurs the highest NSE penalty as it has no knowledge of R_N. RI causes more NSEs, as its reward function does not fully model the penalties for mild and severe NSEs. Overall, the results show that AFS consistently mitigates NSEs, without affecting the task performance substantially.

Table S1: Average cost at task completion.

Method	**Avg. Cost**
Oracle	3759.8 ± 0.00
Naive	61661.0 ± 0.00
RI	71716.6 ± 0.00
AFS (Ours)	1726.5 ± 0.00

2 HUMAN SUBJECTS STUDY IN SIMULATION

2.1 User Interface for Feedback Collection

Figure S2 illustrates the interface used in the simulation-based human subjects study. Participants interacted with the simulated robot through a GUI consisting of feedback buttons whose labels and available options varied depending on the feedback format. For each query, the interface would display the action the agent intends to take in the gridworld and provide a corresponding set of input buttons to record user feedback.

This interface design allowed participants to provide both categorical and comparative feedback efficiently. After an initial training phase to practice providing feedback in different formats, participants self-reported the probability, $\psi(f)$, of providing feedback in a given format f, and effort ratings, $C(f)$.

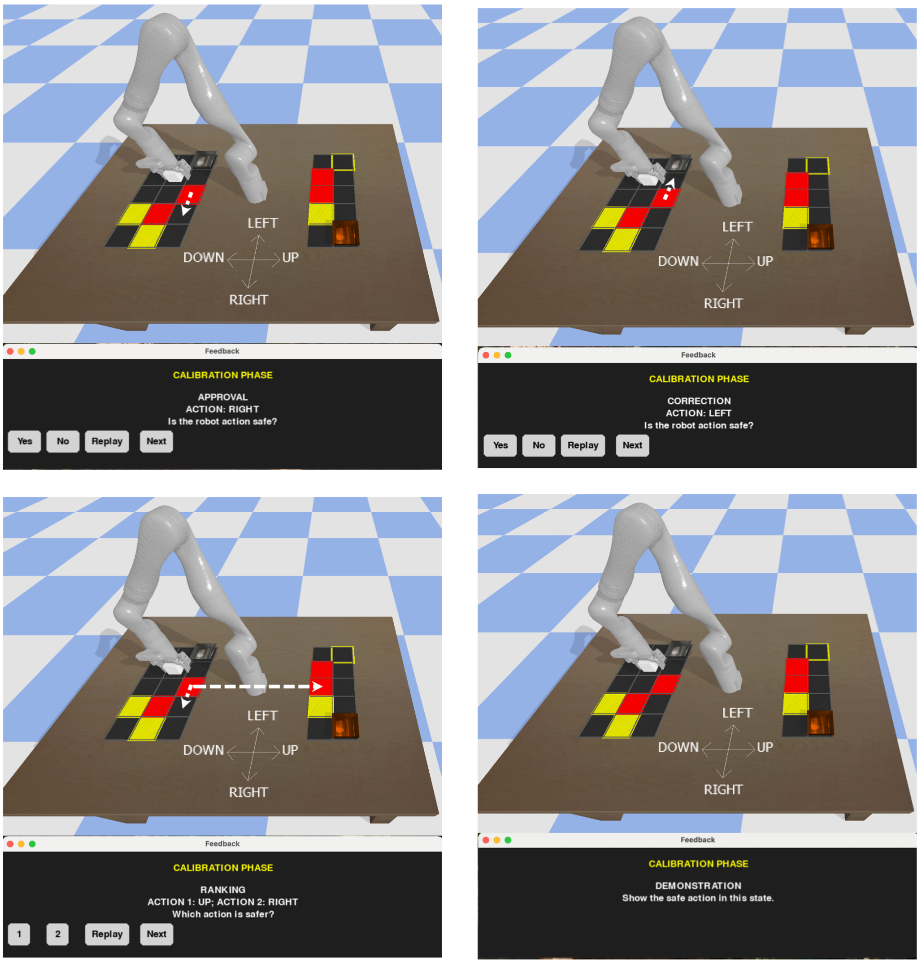


Figure S3: Interface and prompts used in the user study with the Kinova arm. The interface displayed a short clip of the robot's action (indicated by the white dotted arrow) along with a dialog box prompting user feedback. **Top row:** Approval and Correction; **Bottom row:** Ranking and DAM formats.

3 IN-PERSON USER STUDY WITH KINOVA ARM

3.1 User Interface and Feedback Modalities

Figure S3 illustrates the interface and feedback mechanisms used during the in-person user study with the Kinova Gen3 7-DoF robotic arm. Participants interacted with the robot through both a GUI and direct physical manipulation of the arm, depending on the feedback format being queried.

Interface Layout The GUI displayed a simulation of the robot's motion on the tabletop, together with clearly labeled directional arrows (*Up, Down, Left, Right*) corresponding to the robot's possible actions. The lower portion of the interface presented a dialog box, that indicates the current condition, the robot's selected action, and the question or instruction corresponding to the current format. Each query consisted of a brief robot motion in simulation, followed by the corresponding GUI prompt. Participants could replay the motion before submitting their feedback.

REFERENCES

Saran, A., Zhang, R., Short, E. S., and Niekum, S. (2021). Efficiently guiding imitation learning agents with human gaze. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*

Zhang, R., Walshe, C., Liu, Z., Guan, L., Muller, K., Whritner, J., et al. (2020). Atarihead: Atari human eyetracking and demonstration dataset. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. vol. 34