CS 690: Human-Centric Machine Learning

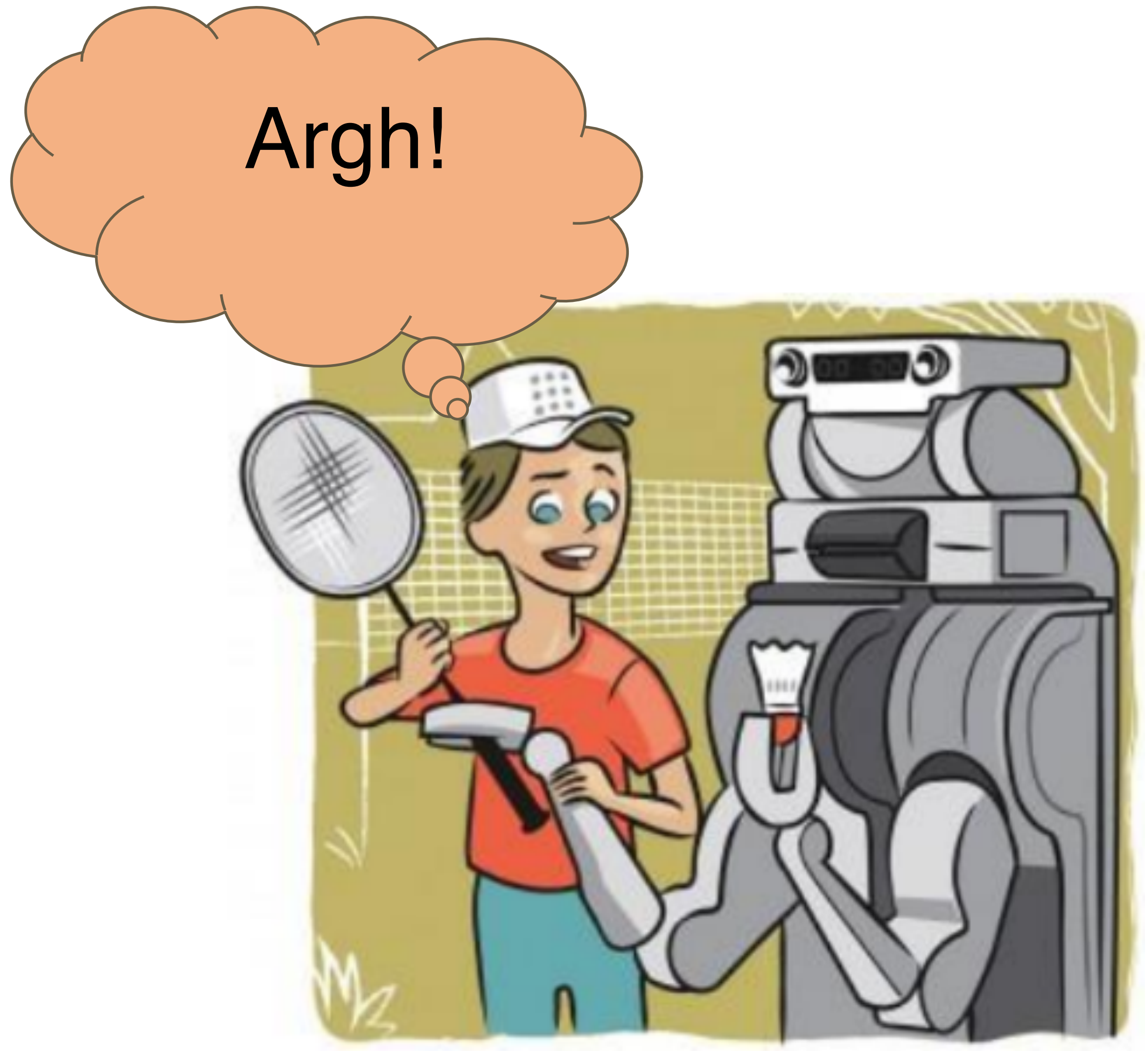
Prof. Scott Niekum

Multimodal reward inference

Slide credits: Akanksha Saran and Yuchen Chi

Humans Leverage more than State-Action Pairs when Learning from one another

Types of Human Social Cues



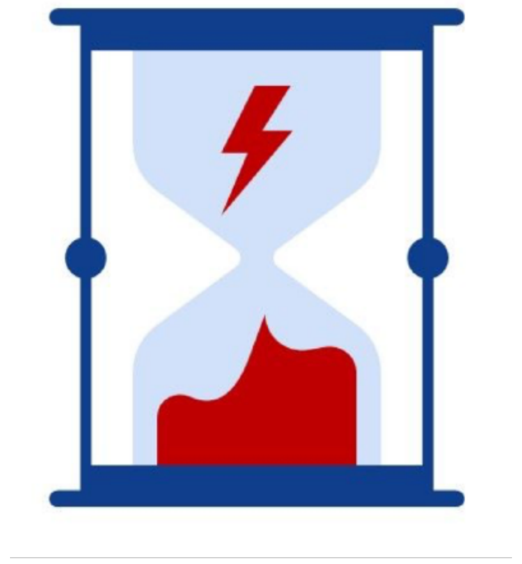
Audio

Eye Gaze

Gestures and Body Pose

Head Pose, Facial Expressions

Key Challenges of Leveraging Human Cues for Learning

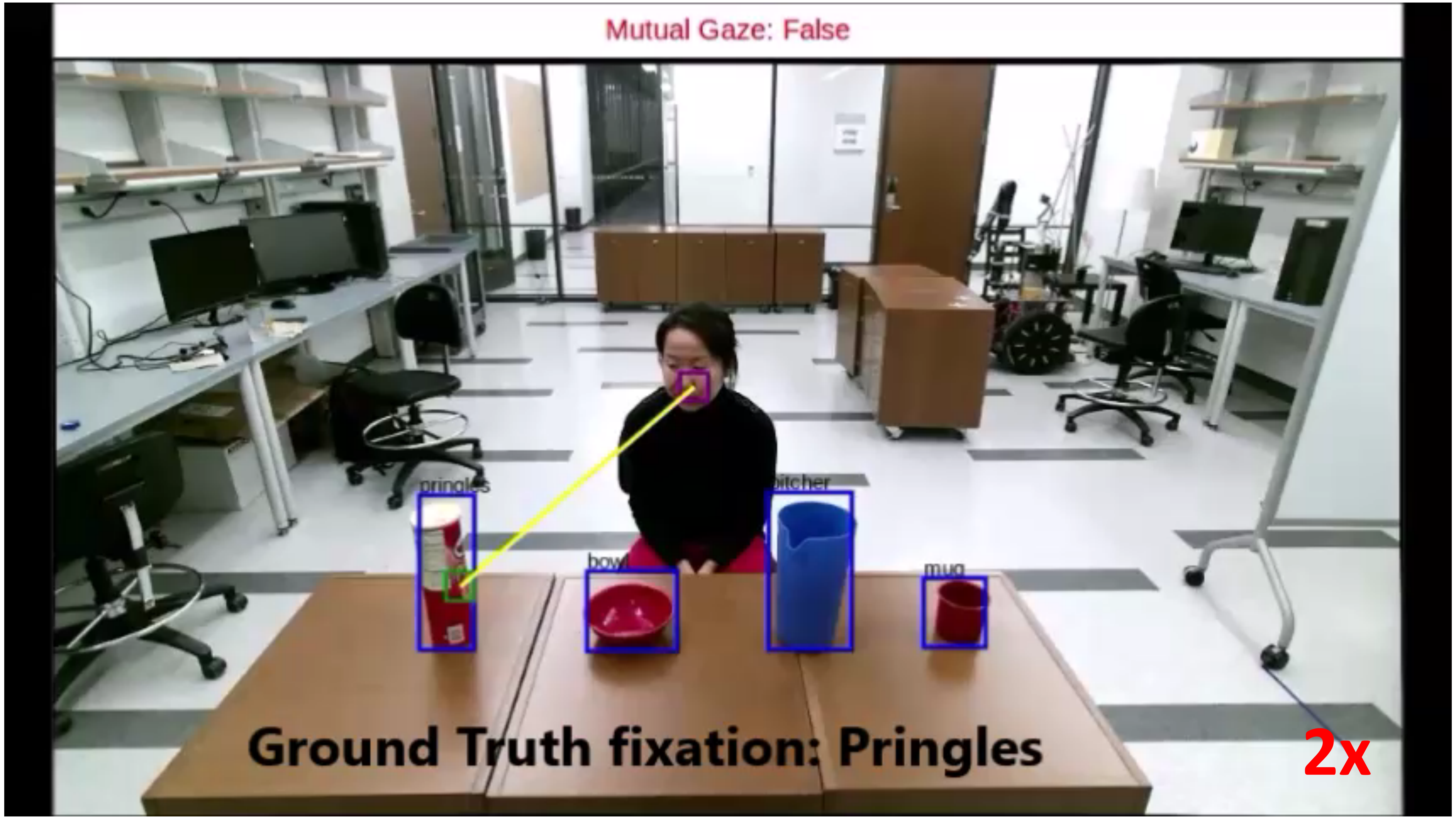


- Demonstrations are expensive to collect
- Human data often only be present at train time
- Humans provide a rich source of behavioral information, however collecting human data requires iterative design and testing

Human Gaze reveals Intentions

Argyle, M. Non-verbal communication in human social interaction. 1972.
Hayhoe, M & Ballard, D. Eye movements in natural behavior. Trends in cognitive sciences, 9(4), 2005.

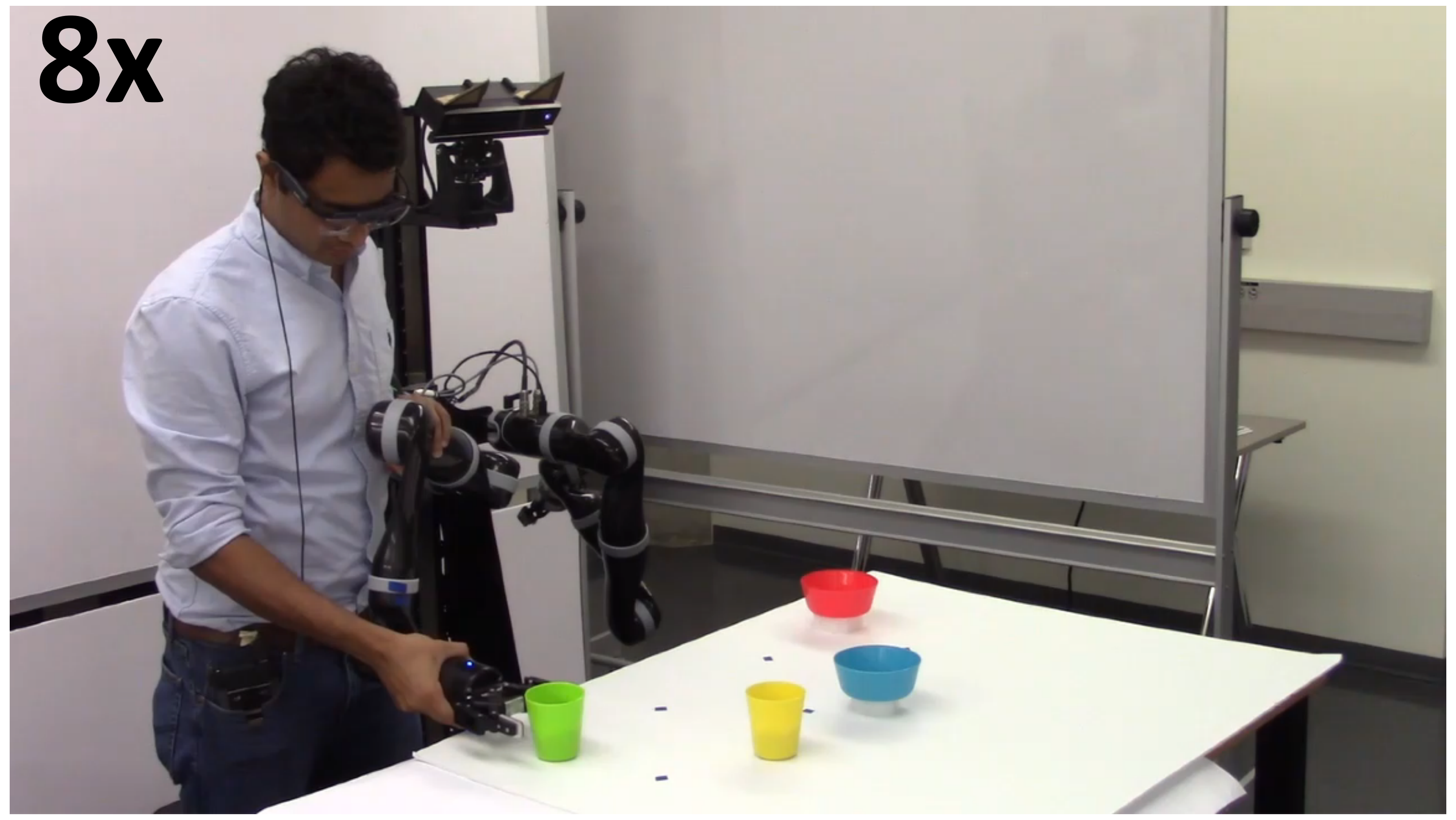
Can a Situated Robot detect Human Gaze Fixations without additional Eye-tracking Hardware?



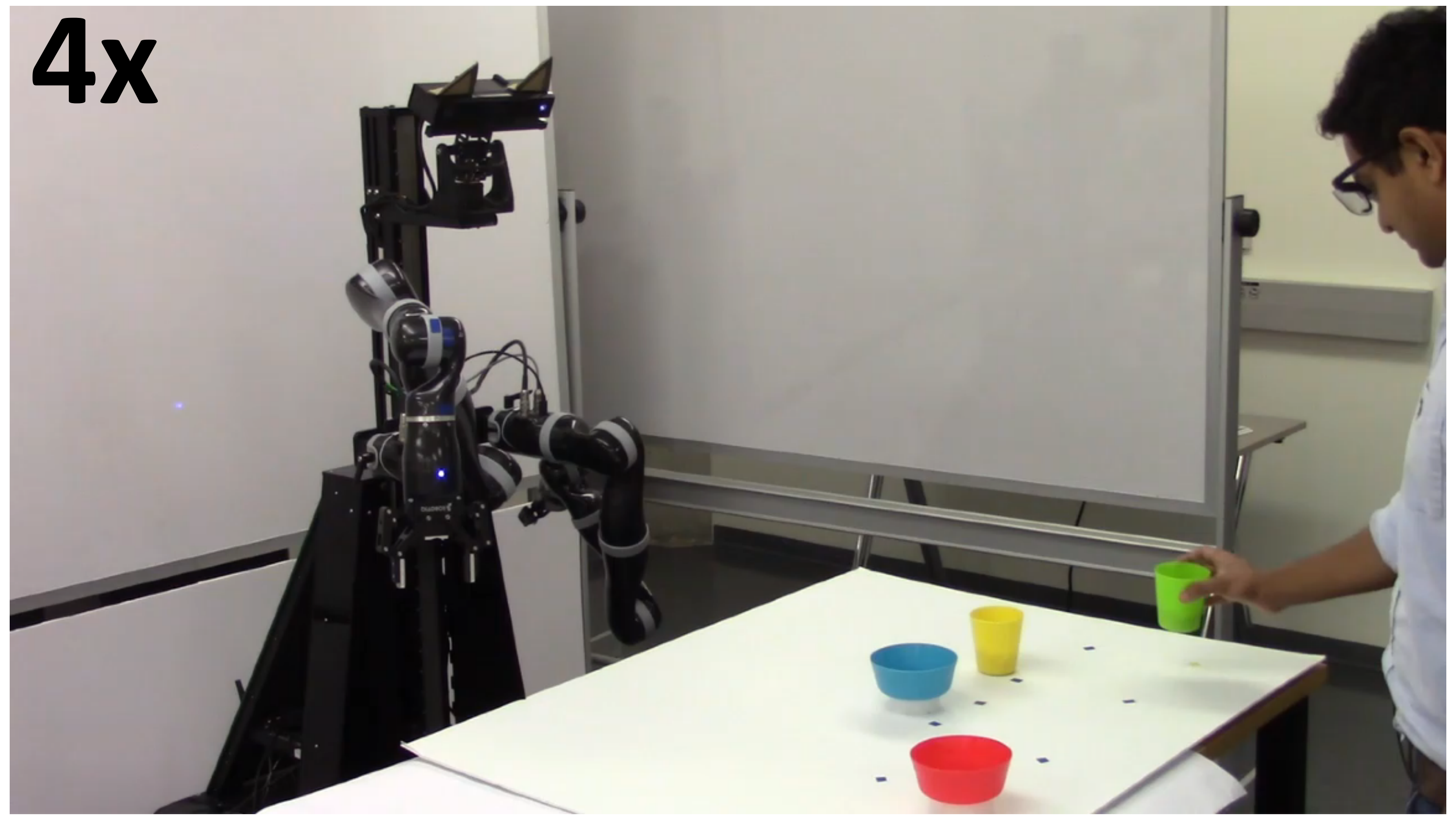
Saran, A., S. Majumdar, E. S. Short, A. Thomaz, and S. Niekum. Real-time Human Gaze Following for Human-Robot Interaction. IROS, 2018.

Gaze Patterns in Human Demonstrations for Robots

Keyframe-based Kinesthetic Teaching (KT)



Observational/Video Demonstrations



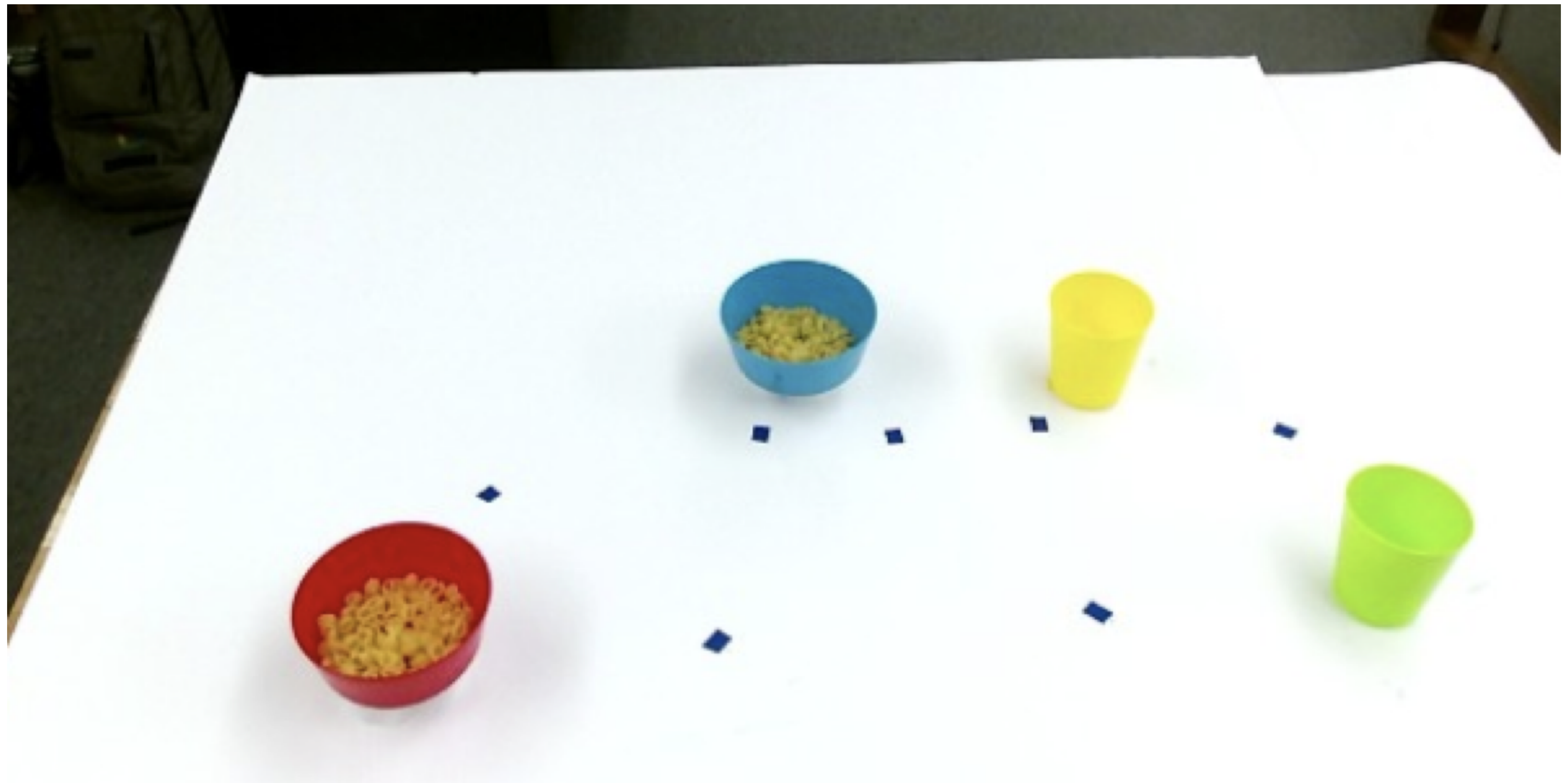
User Study and Data Collection

- Tobii Pro Glasses 2 Eye Tracker
- 20 subjects:
  - 10 expert robot users
  - 10 novice robot users
- Demonstration Types:
  - Kinesthetic Demonstrations (~124 mins)
  - Video Demonstrations (~27 mins)
- Tasks:
  - Placement (single-step)
  - Pouring (multi-step)

Saran, A., E. S. Short, A. Thomaz, and S. Niekum. Understanding teacher gaze patterns for robot learning. CoRL, 2019.

Understanding Human Gaze of Demonstrators for Embodied Robots

Pouring Task



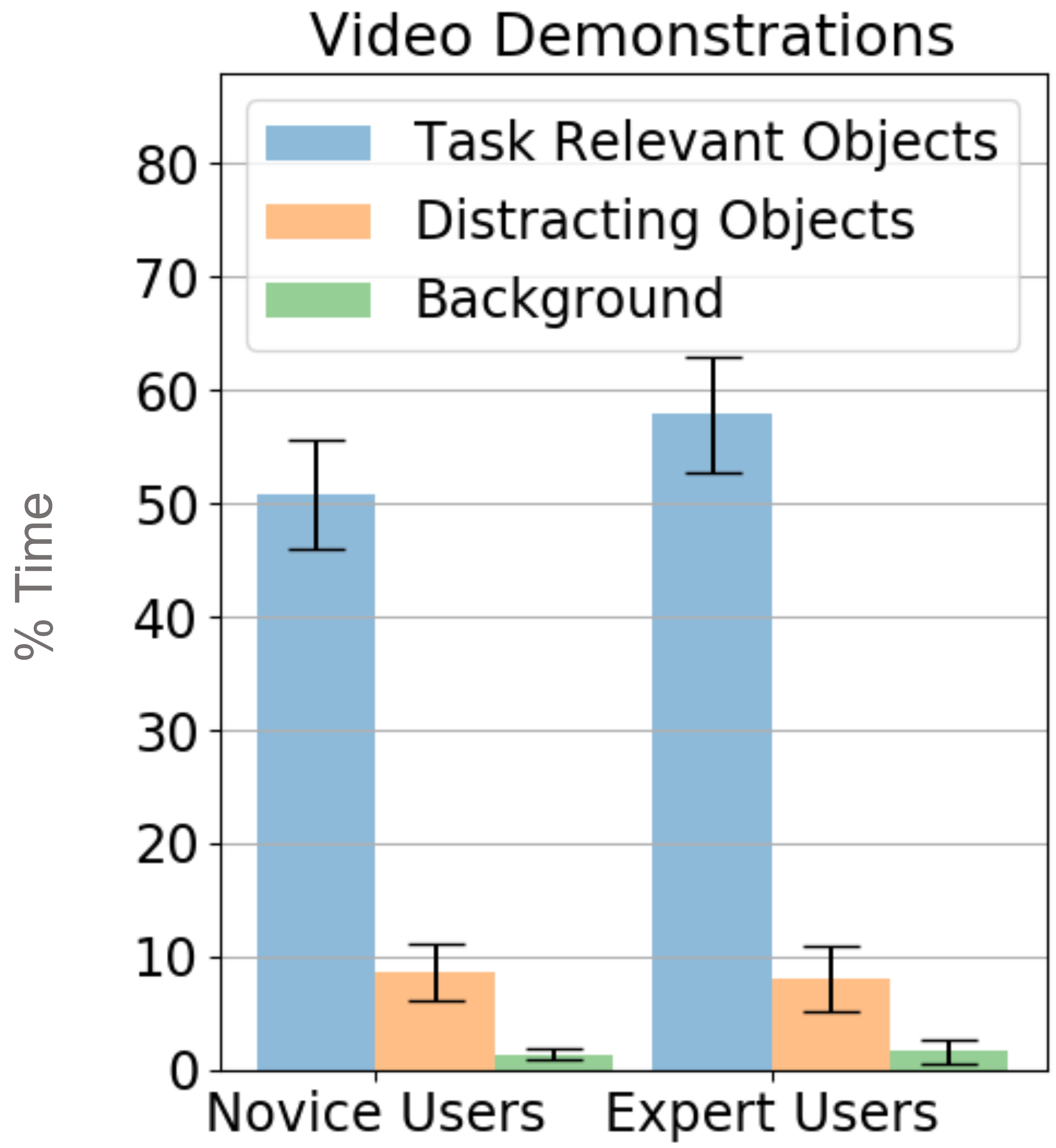
"Pour pasta from green cup into red bowl and from yellow cup into blue bowl"

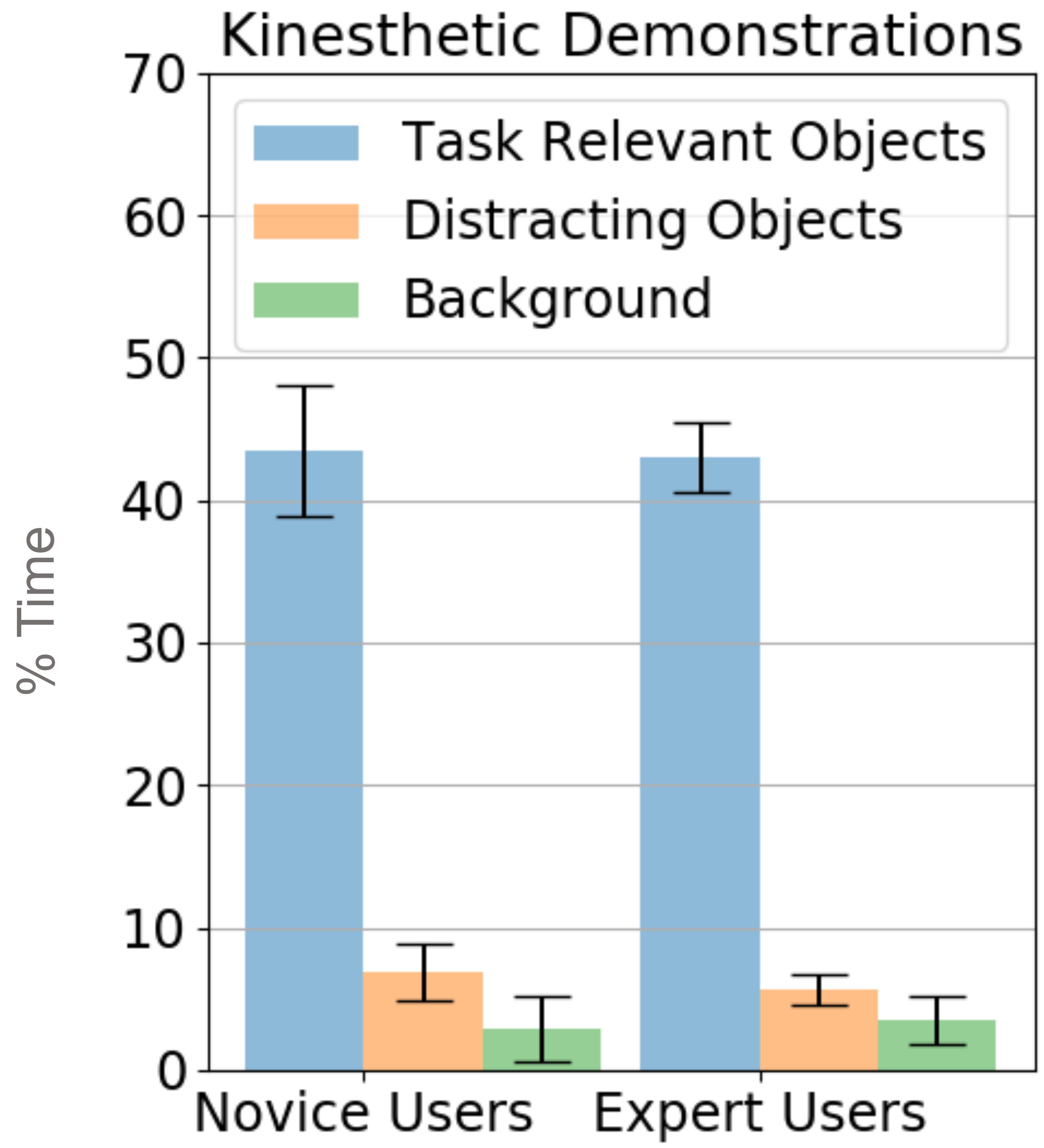
Placement Task

"Place the green ladle to the left of red plate"
"Place the green ladle to the right of yellow bowl"

Saran, A., E. S. Short, A. Thomaz, and S. Niekum. Understanding teacher gaze patterns for robot learning. CoRL, 2019.

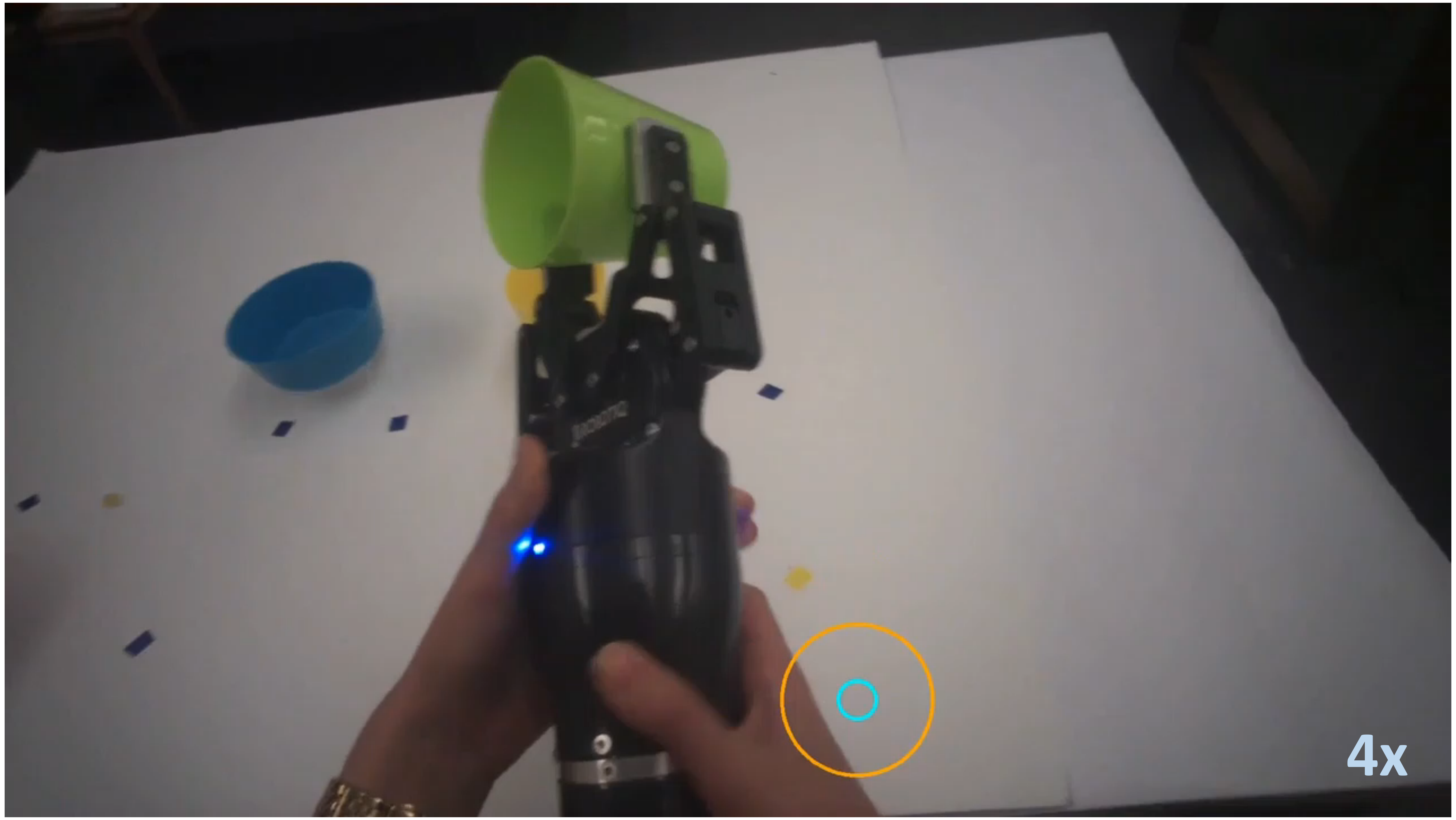
Video and Kinesthetic Demos: Users focus their Gaze on Task-Relevant objects

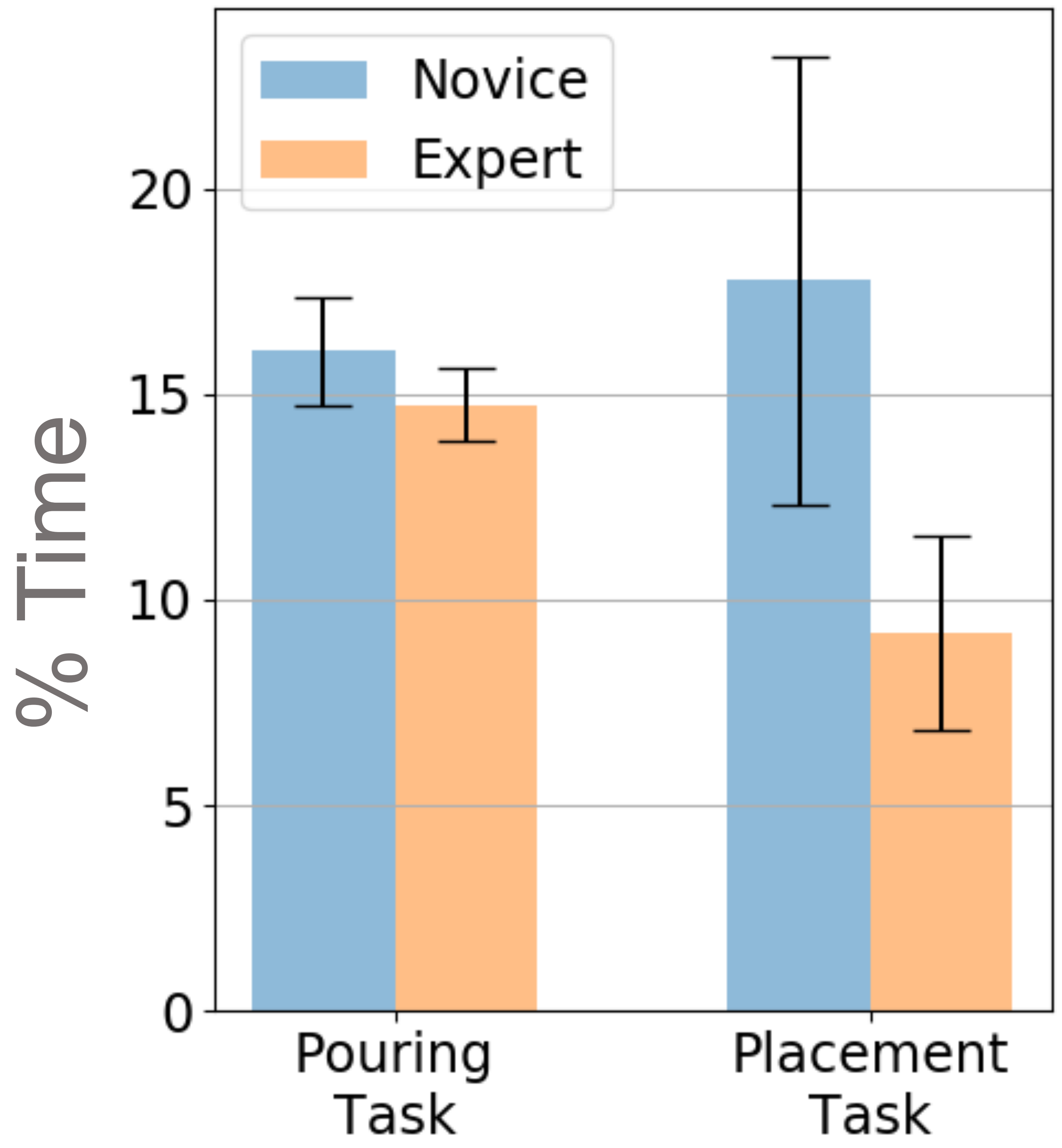




Saran, A., E. S. Short, A. Thomaz, and S. Niekum. Understanding teacher gaze patterns for robot learning. CoRL, 2019.

Kinesthetic Demos: Novice Users focus more on the Robot's Gripper





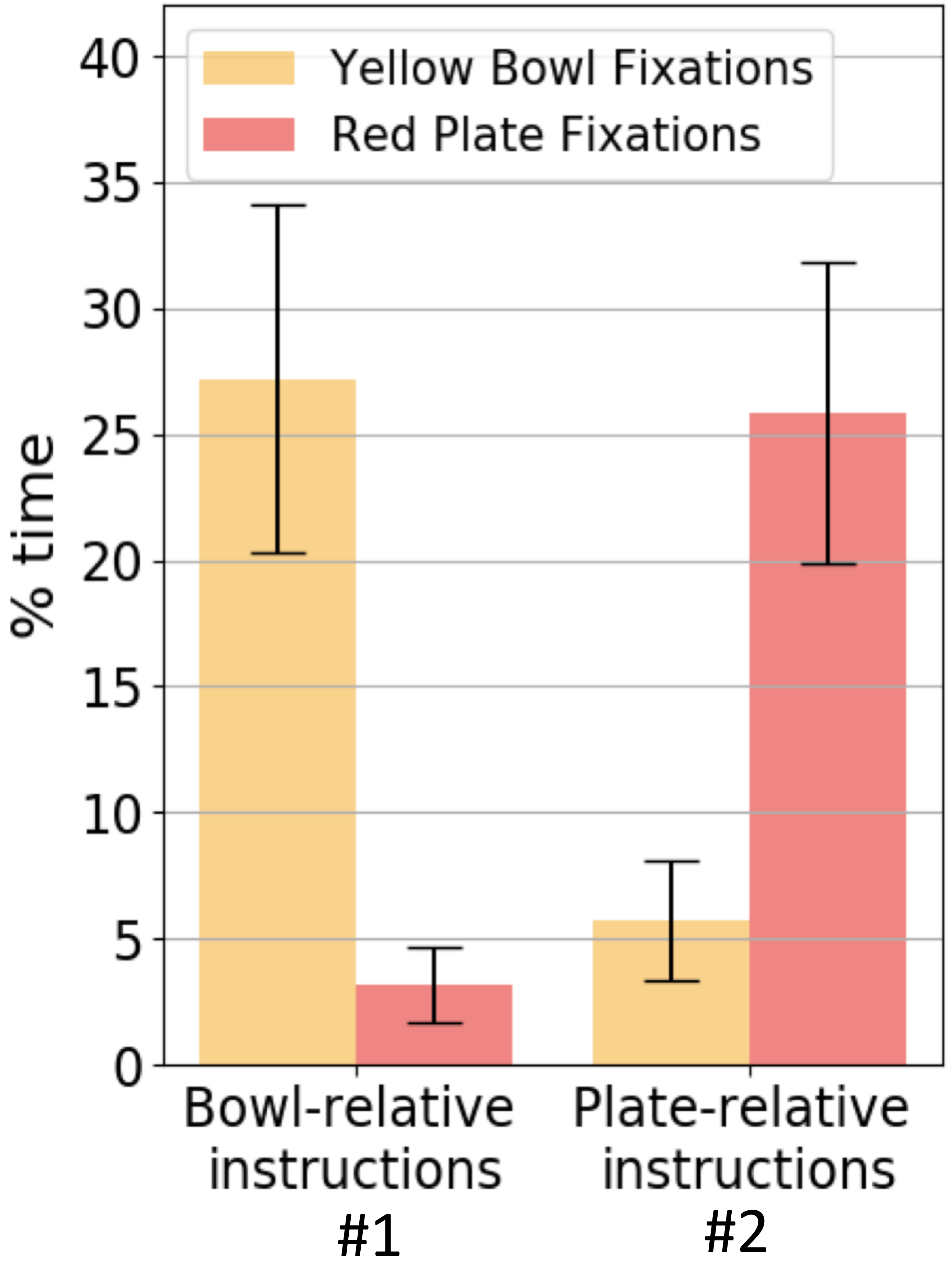
Saran, A., E. S. Short, A. Thomaz, and S. Niekum. Understanding teacher gaze patterns for robot learning. CoRL, 2019.

Most Gaze Fixations are on Objects of Interest under Ambiguous Demos

Two instructions:

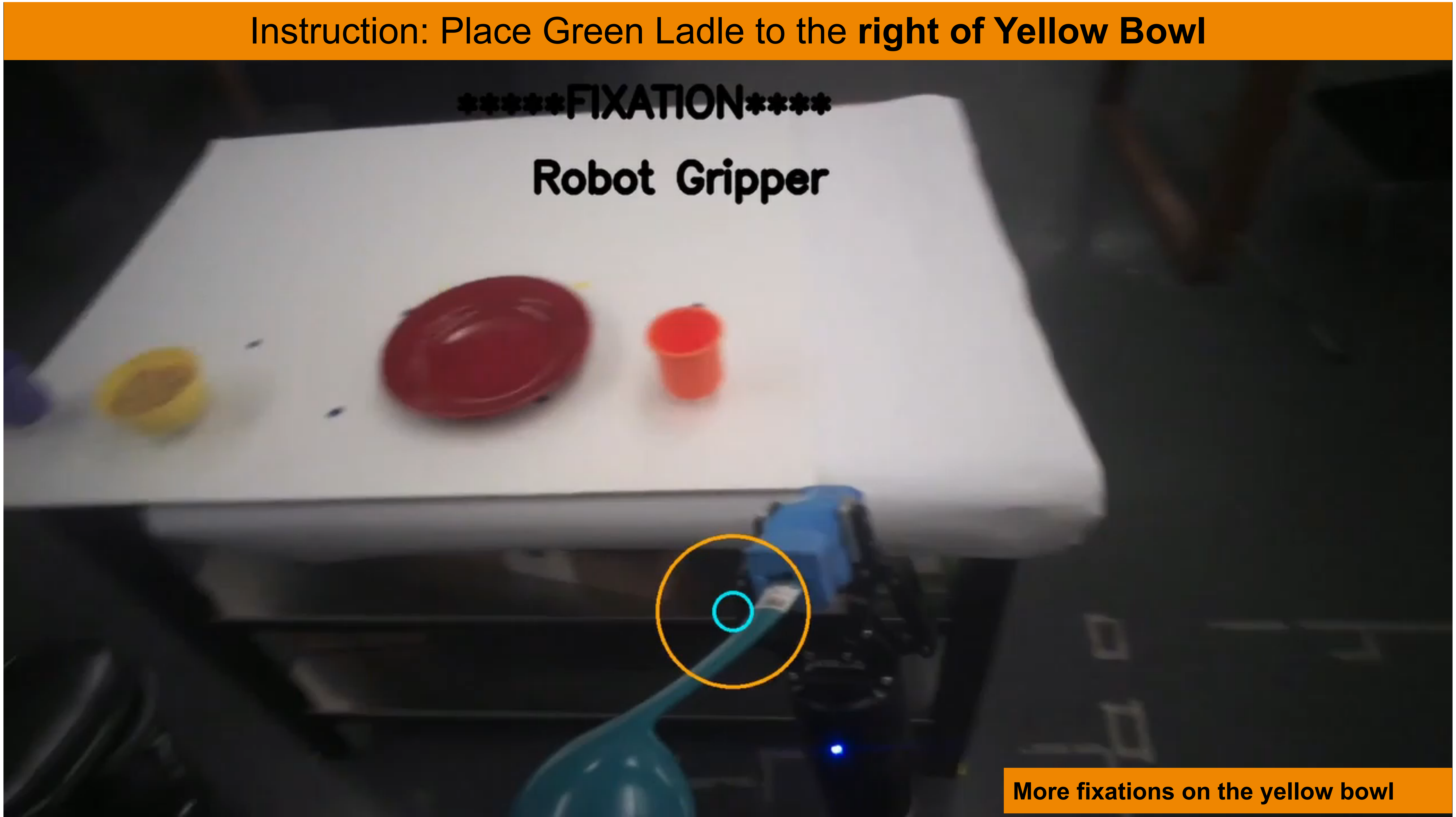
1: “Place the green ladle to the right of yellow bowl”

2: “Place the green ladle to the left of red plate”

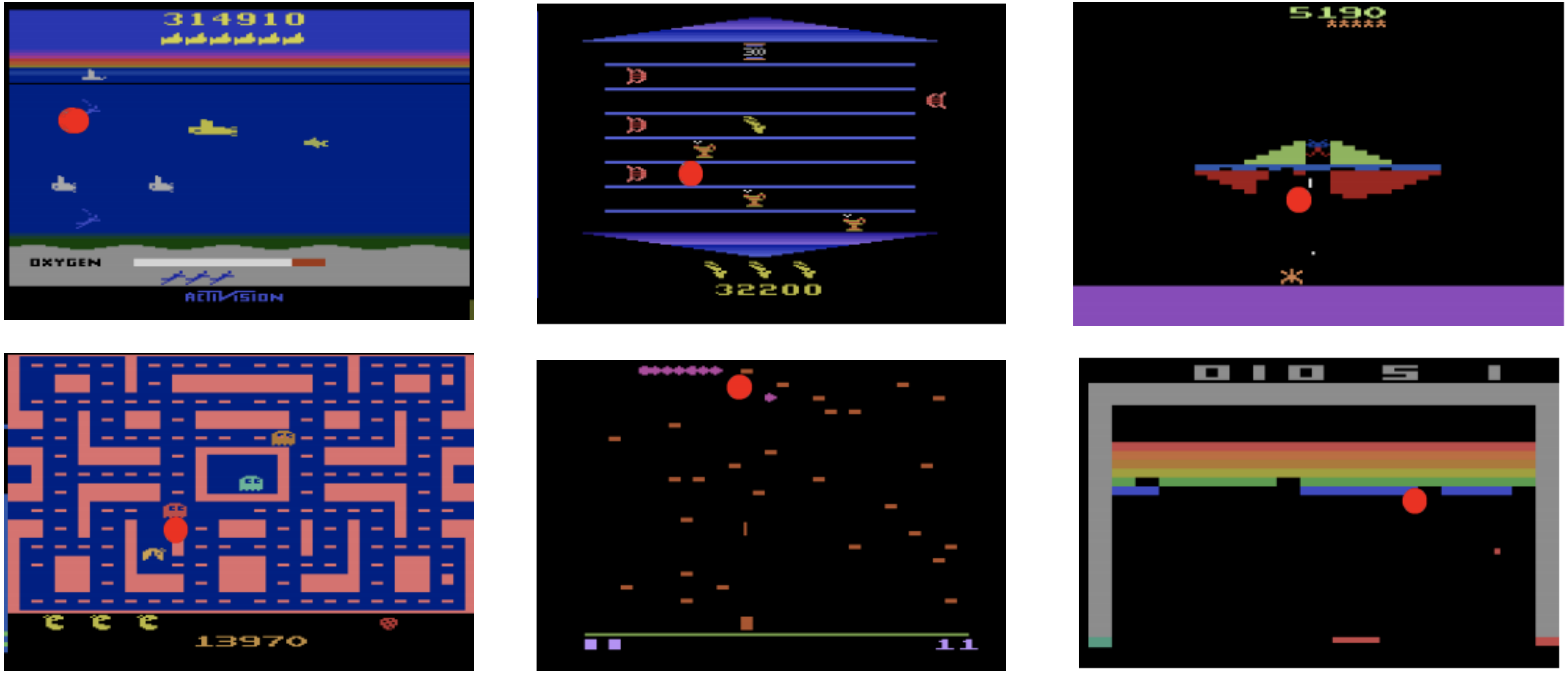


Saran, A., E. S. Short, A. Thomaz, and S. Niekum. Understanding teacher gaze patterns for robot learning. CoRL, 2019.

Gaze Fixations during Ambiguous Placement Demonstrations

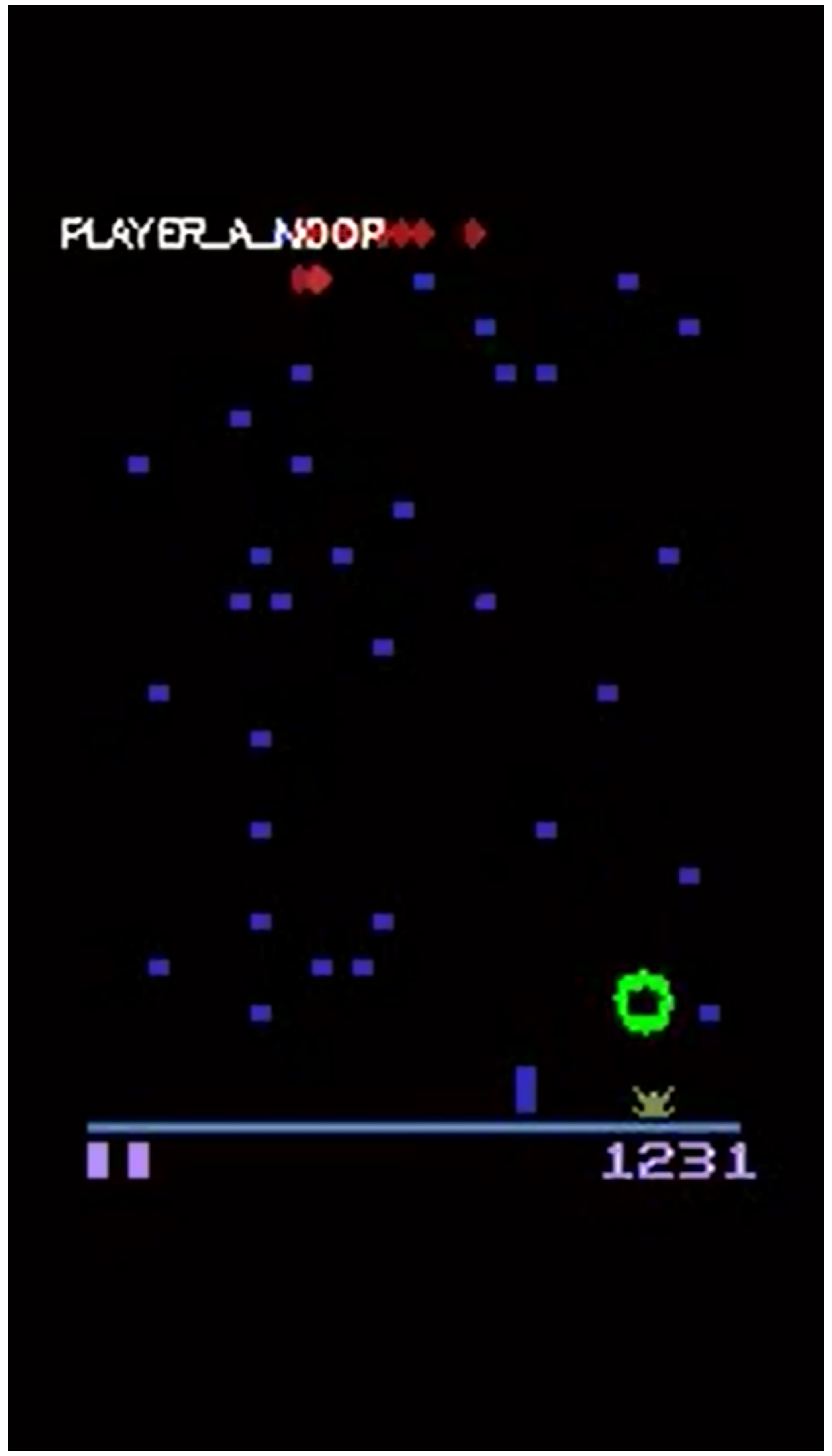


Analyzing Human Gaze of Demonstrators for Simulated Agents



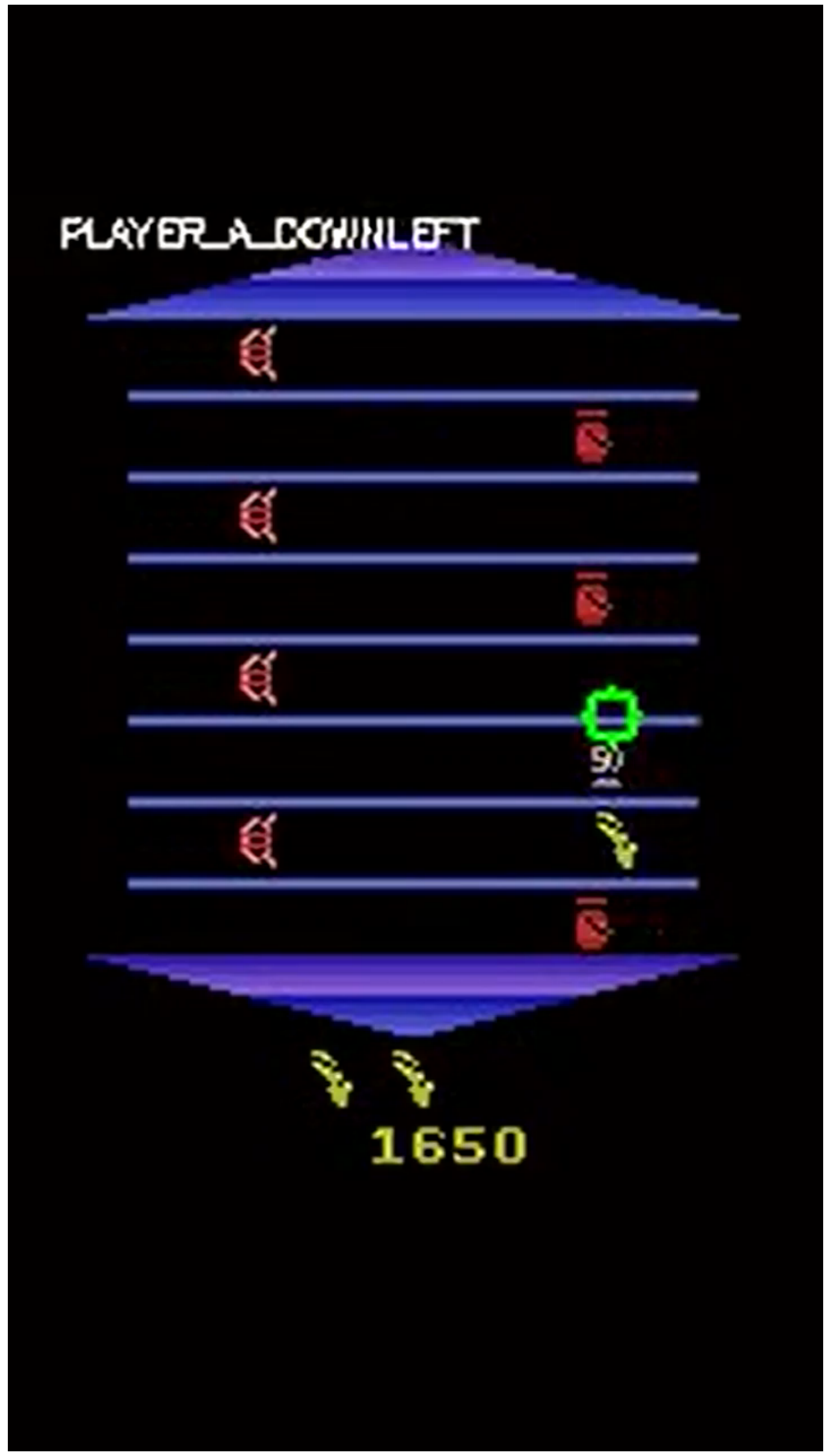
Attention on Objects of Interest for the next Action

Centipede



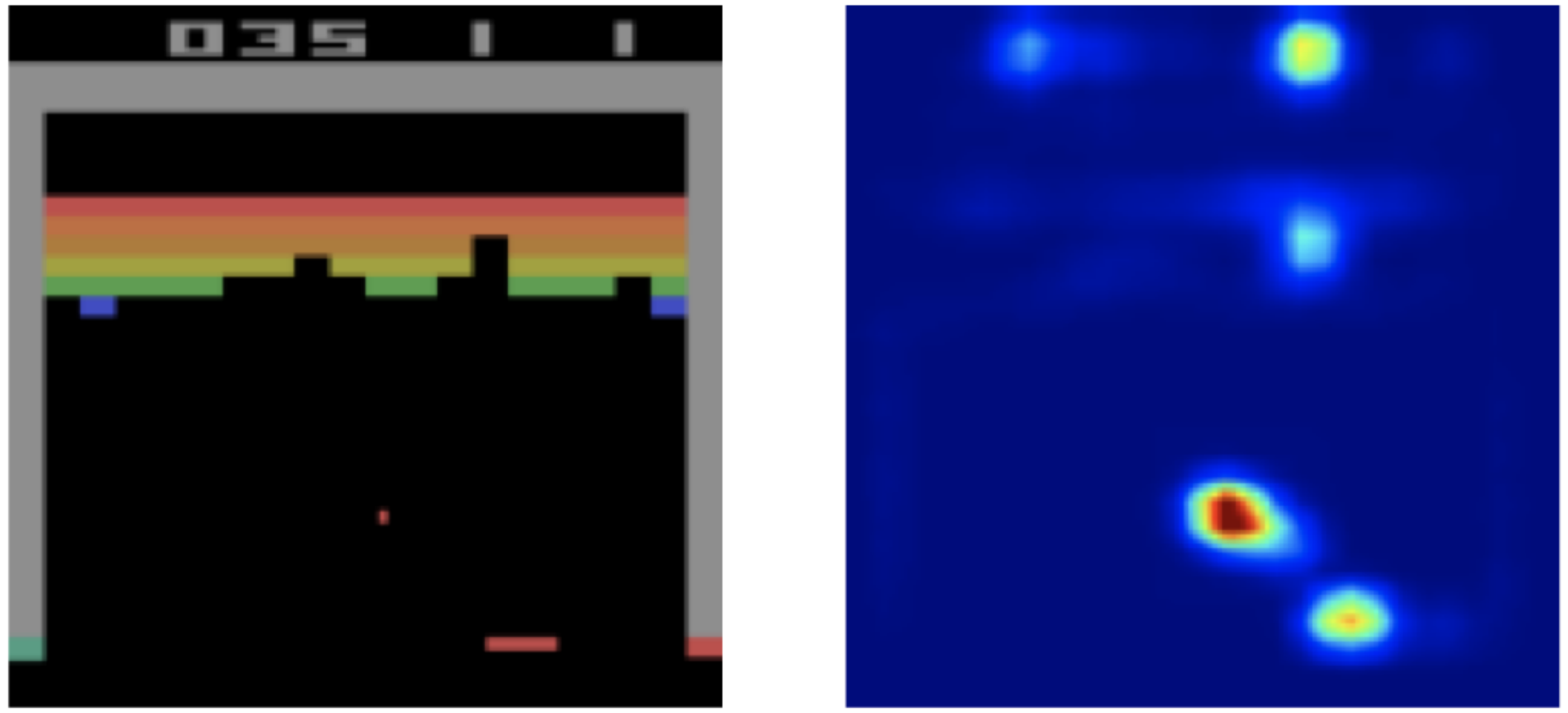
Gaze indicates where the human might shoot next

Asterix



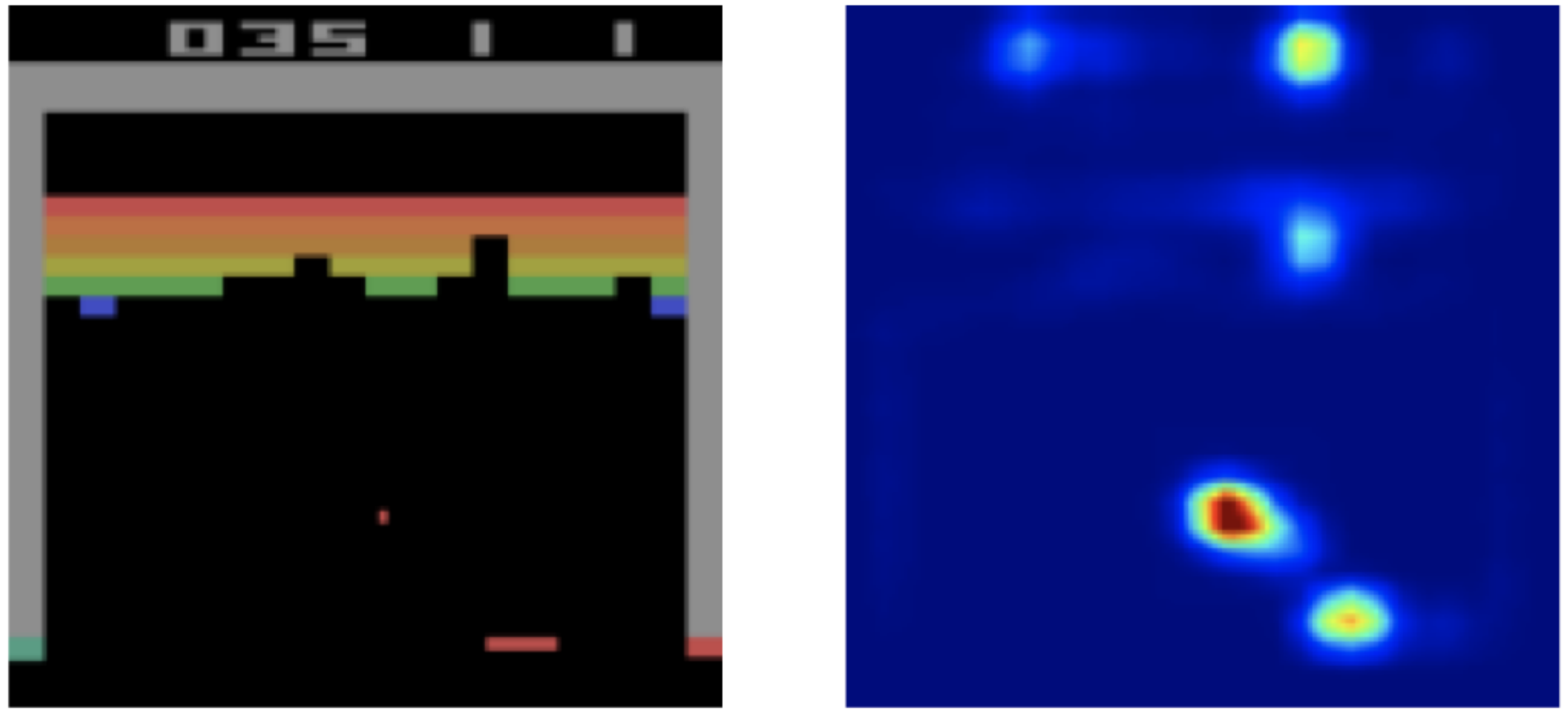
Gaze on food that should be eaten and dynamite which should be avoided

What do RL agents attend to?



(a) Game State (b) RL Attention

What do RL agents attend to?



(a) Game State (b) RL Attention

Perturbation based method to compute RL attention

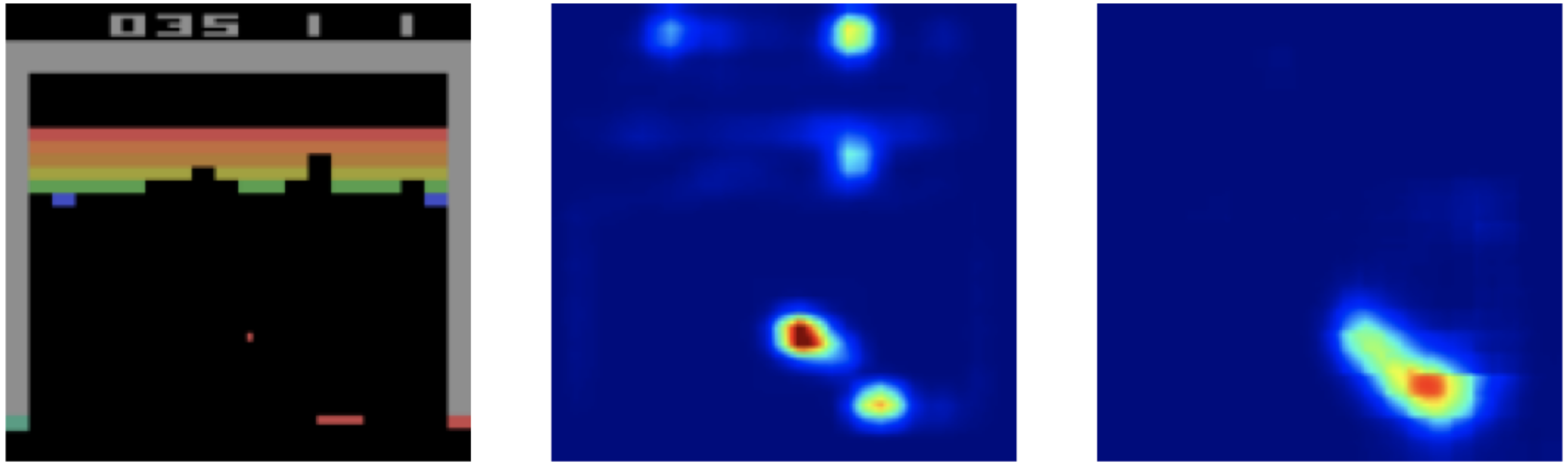
$$S_\pi(i,j) = \frac{1}{2}||\pi(I) - \pi(\phi(I,i,j))||^2$$

Change in policy by perturbing the image at a pixel

Greydanus, S., Koul, A., Dodge, J., & Fern, A. Visualizing and understanding Atari agents. ICML, 2018.

RL agent attention “covers” regions attended by human gaze

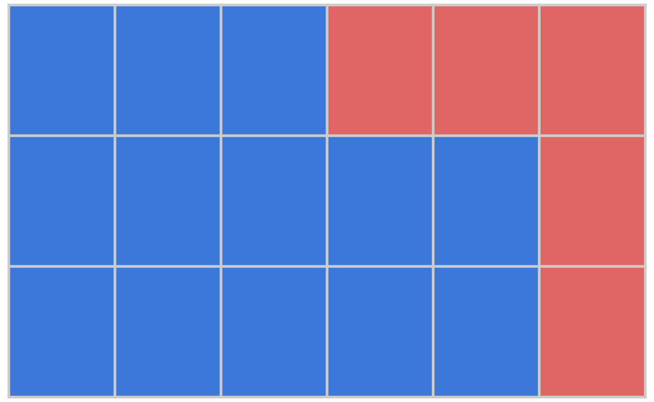
… while also attending to other regions



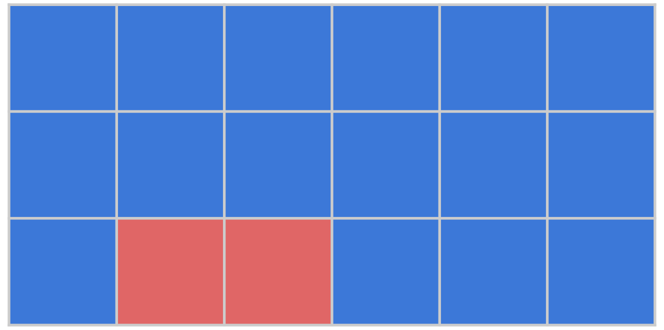
(a) Game State (b) RL Attention (c) Human Gaze

Coverage Metric

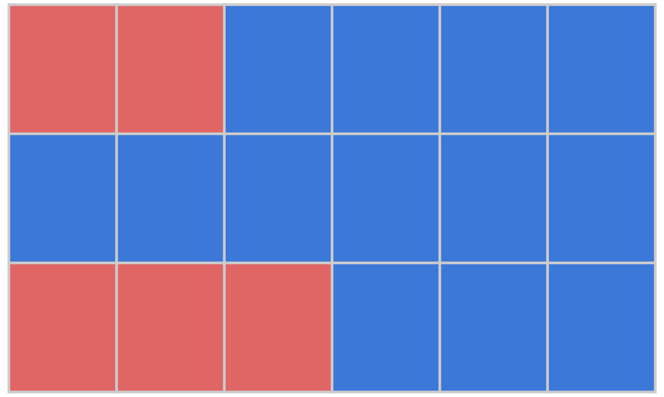
$$KL(P||Q) = \sum_i \sum_j P(i,j) \log \left(\frac{P(i,j) + \epsilon}{Q(i,j) + \epsilon} \right)$$



Q: RL Attention Map (**No Coverage**)
KL(P || Q) = **8.5**



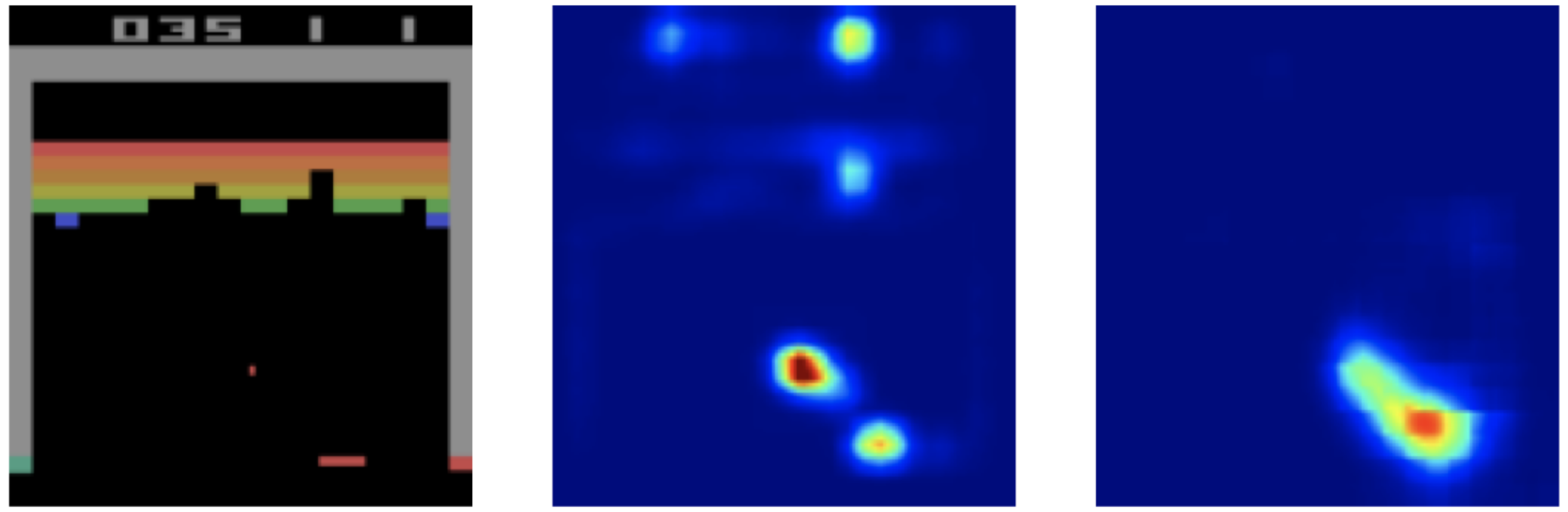
P: Human Gaze map



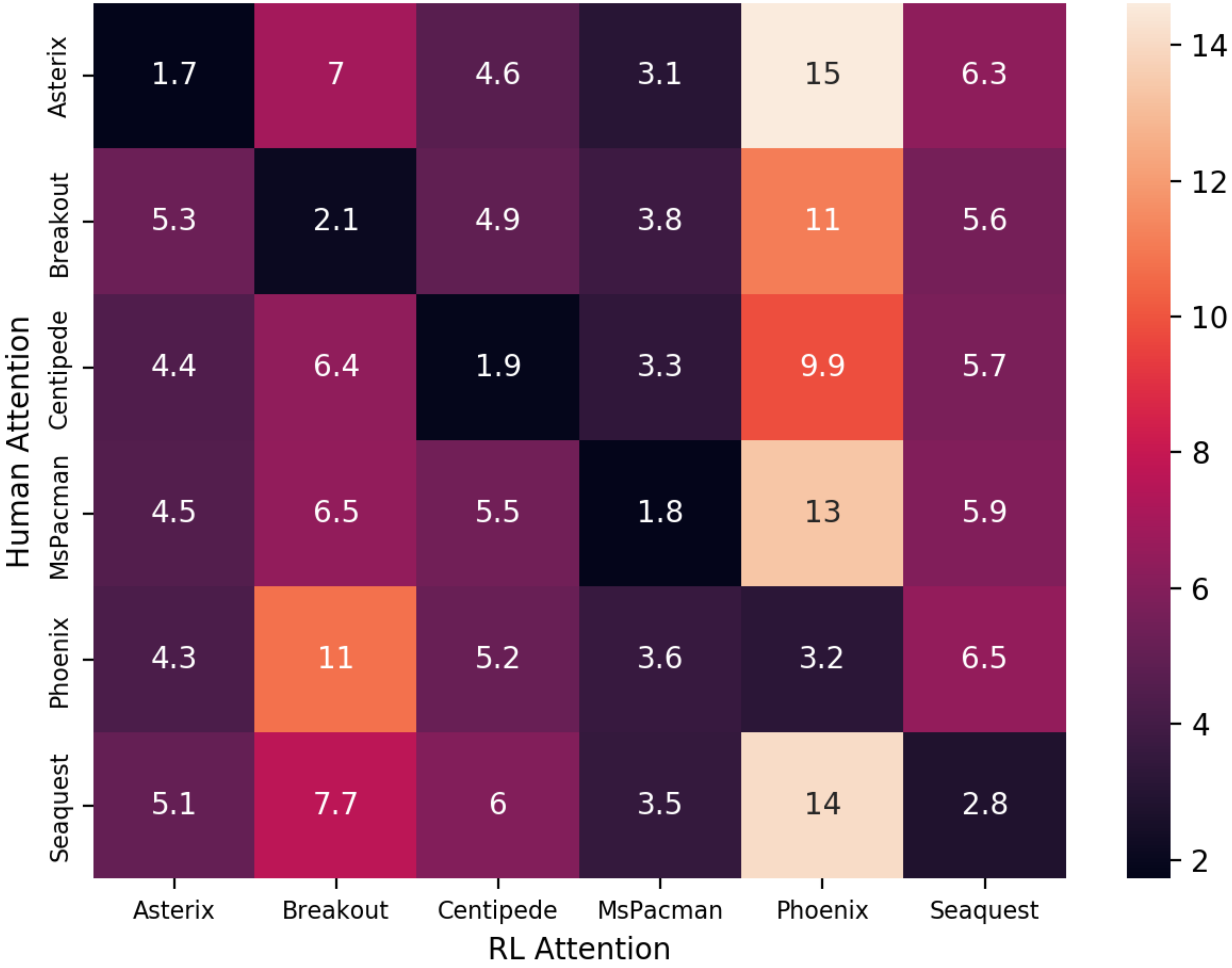
Q: RL Attention Map (**Has Coverage**)
KL(P || Q) = **0.9**

Bylinskii, Z., T. Judd, A. Oliva, A. Torralba, and F. Durand. “What do different evaluation metrics tell us about saliency models?”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.

Comparison of Human Attention and RL agent Attention



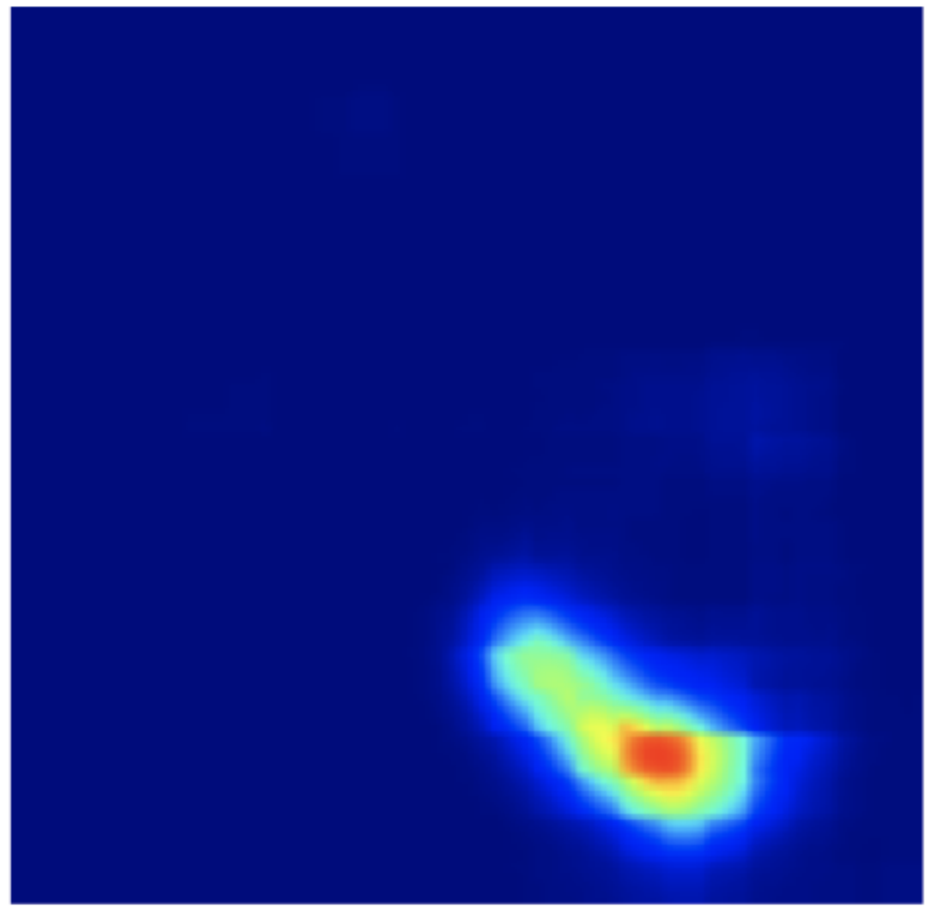
(a) Game State (b) RL Attention (c) Human Gaze



Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

How to effectively leverage gaze for imitation learning?
What if gaze is only available at train time?

- Prior approaches use gaze as an input required at test time
- Need to model gaze data per task
- A simpler alternative – guide the training based on gaze data available at train time



Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

Gaze as a supervisory signal for existing imitation learning methods

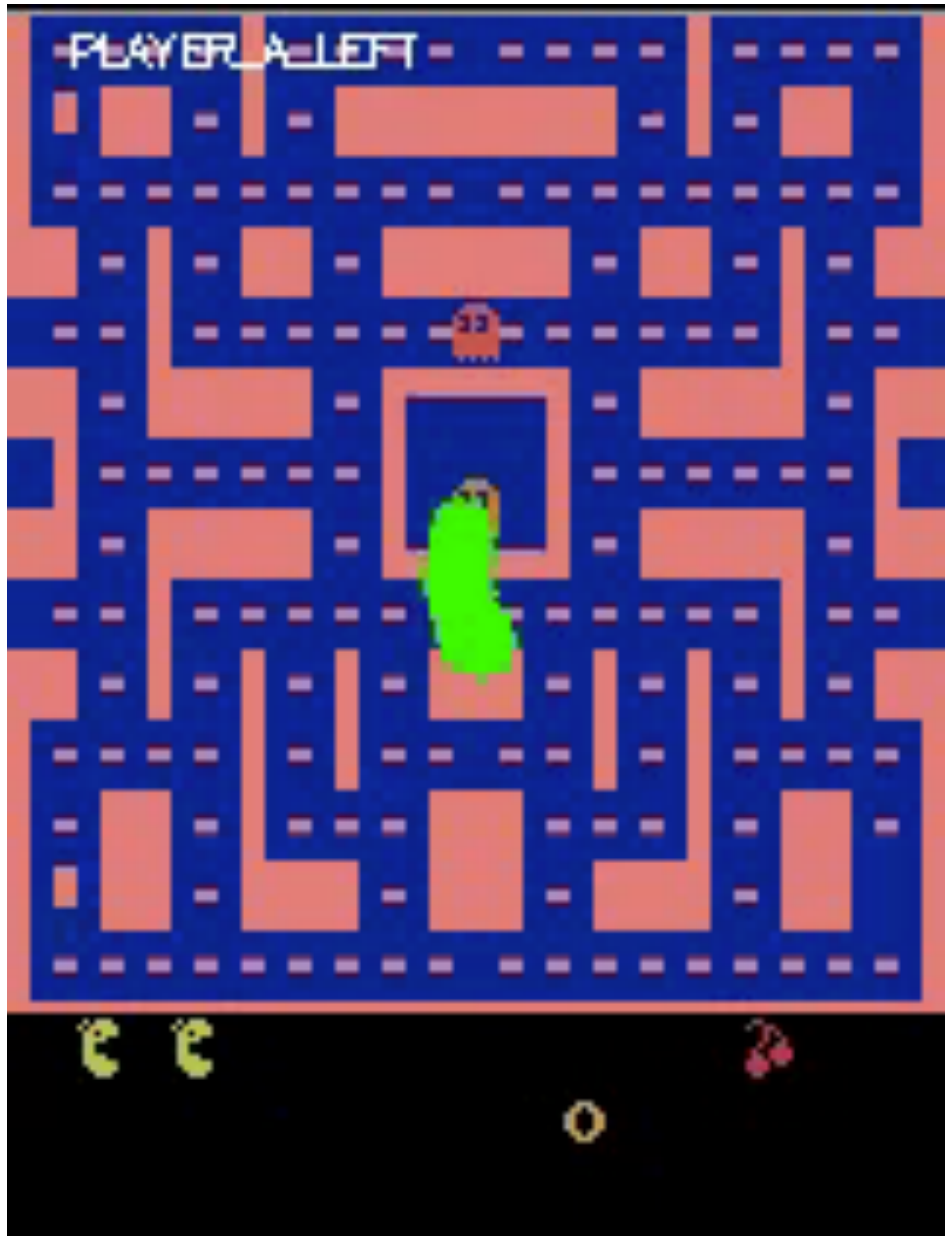
Use an auxiliary coverage-based gaze loss **(CGL)** to guide the attention of existing imitation learning methods

- Three Imitation Learning methods: BC, BCO, TREX
- 20 Atari games with varying complexity, dynamics, visual features and rewards
- Compare with prior state-of-the-art gaze-augmentation LfD methods

Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

Atari-HEAD: Atari Human Demonstrations and Gaze Dataset

- Human gaze and demonstration data for 20 Atari Games
- EyeLink 1000 eye tracker at 1000Hz
- Total data worth 117 hours collected with 4 users



Zhang, R., C. Walshe, Z. Liu, L. Guan, K. S. Muller, J. A. Whritner, L. Zhang, M. M. Hayhoe, and D. H. Ballard. "Atari-head: Atari human eye-tracking and demonstration dataset." AAAI, 2020.

CGL loss

$$CGL(g, f^{'}) = \sum_{i \in (1,h)} \sum_{j \in (1,w)} g_{i,j} \left[\log \frac{g_{i,j} + \epsilon}{f^{'}_{i,j} + \epsilon} \right]$$

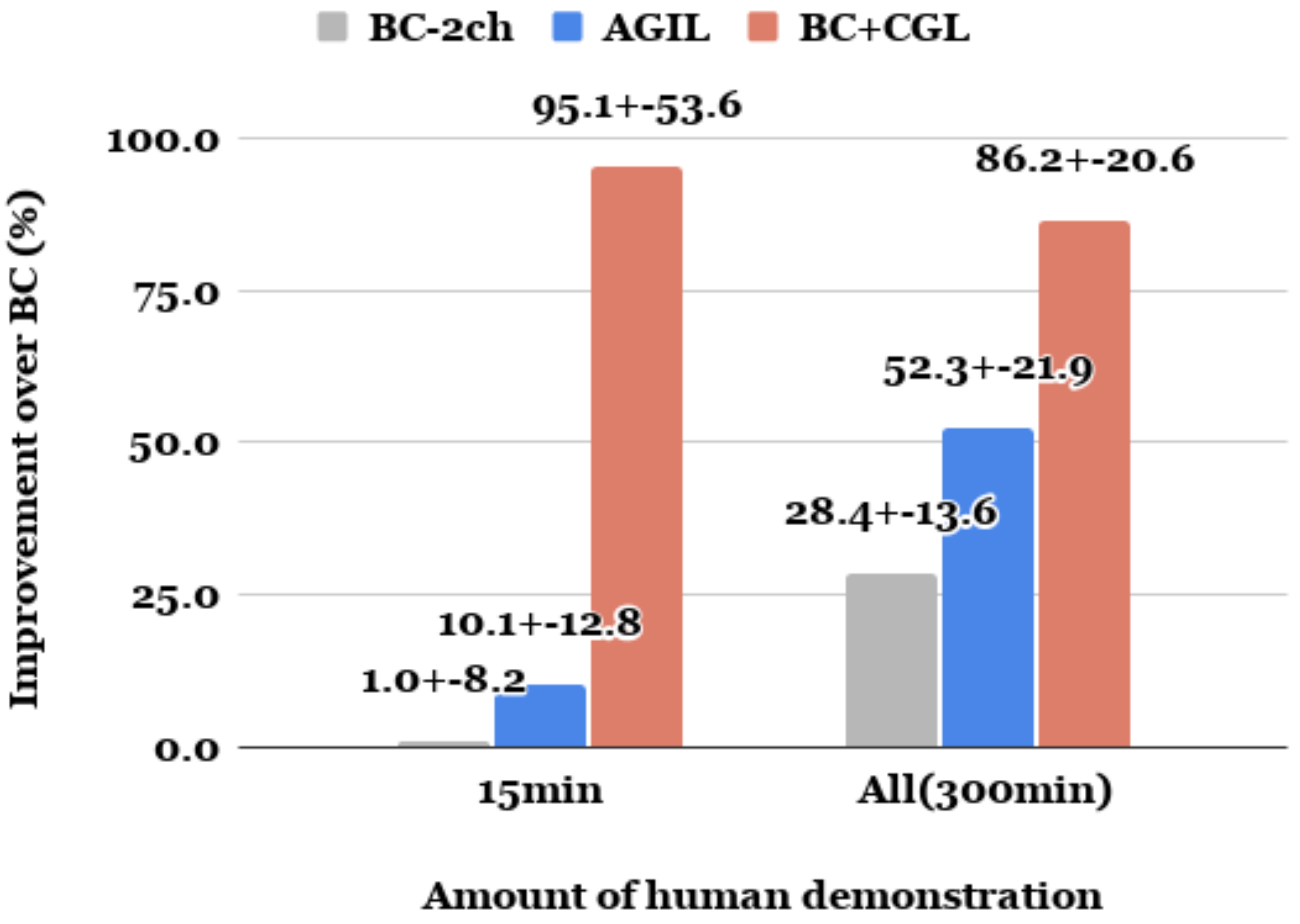
where

$$f^{'}_{i,j} = \frac{\exp^{f_{i,j}}}{\sum_{k=0}^{k=h-1} \sum_{j=0}^{j=w-1} \exp^{f_{k,l}}}$$

CGL improves performance for 3 imitation learning algorithms

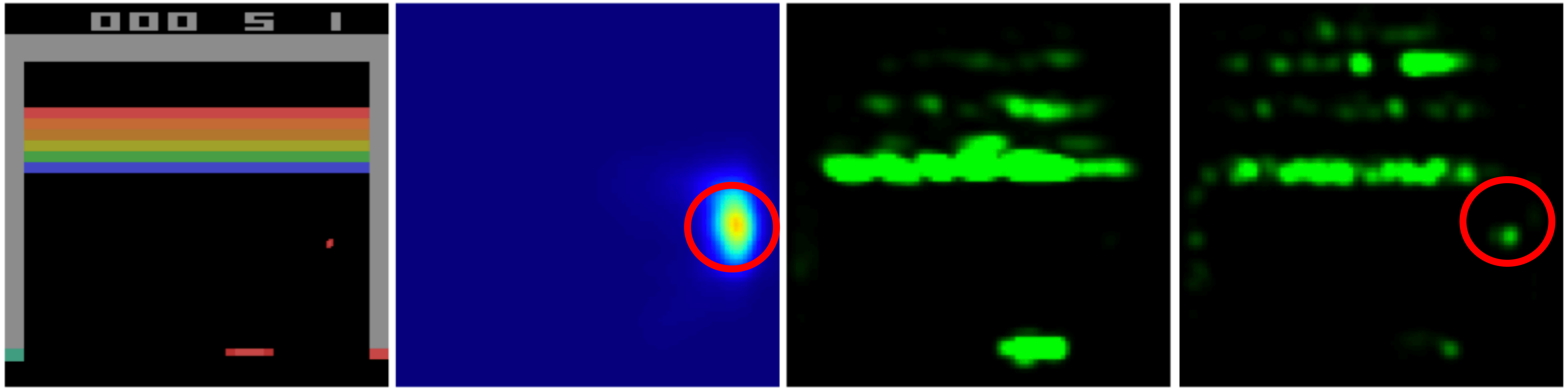
IL Algorithm	% Improvement with CGL
BC	160%
BCO	343%
TREX	390%

CGL outperforms existing Gaze-augmentation methods for Imitation Learning



Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

CGL Agents attend to Visual Features from Human's Overt Attention

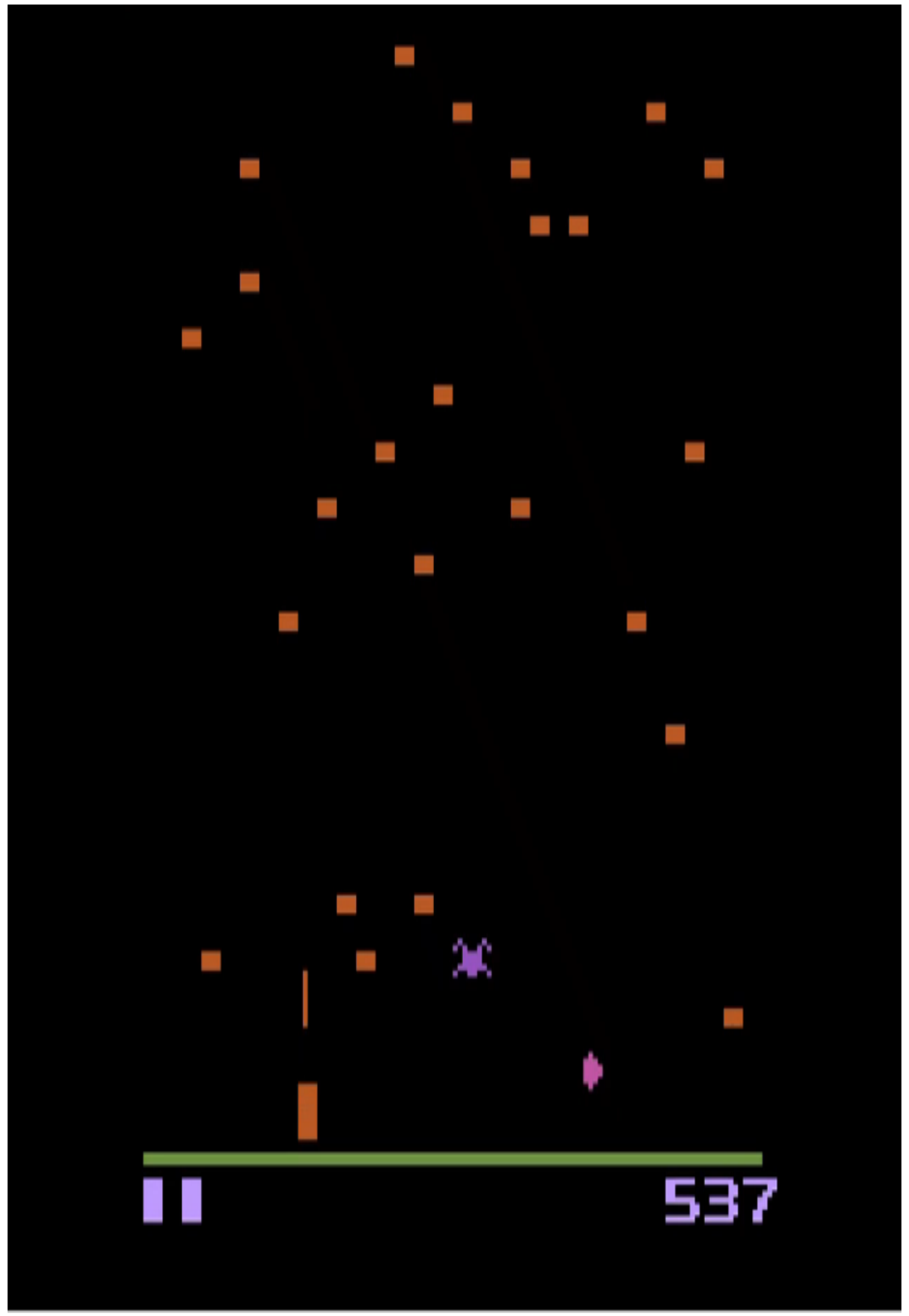


(a) Input image (b) Human (c) T-REX (d) T-REX+CGL

Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

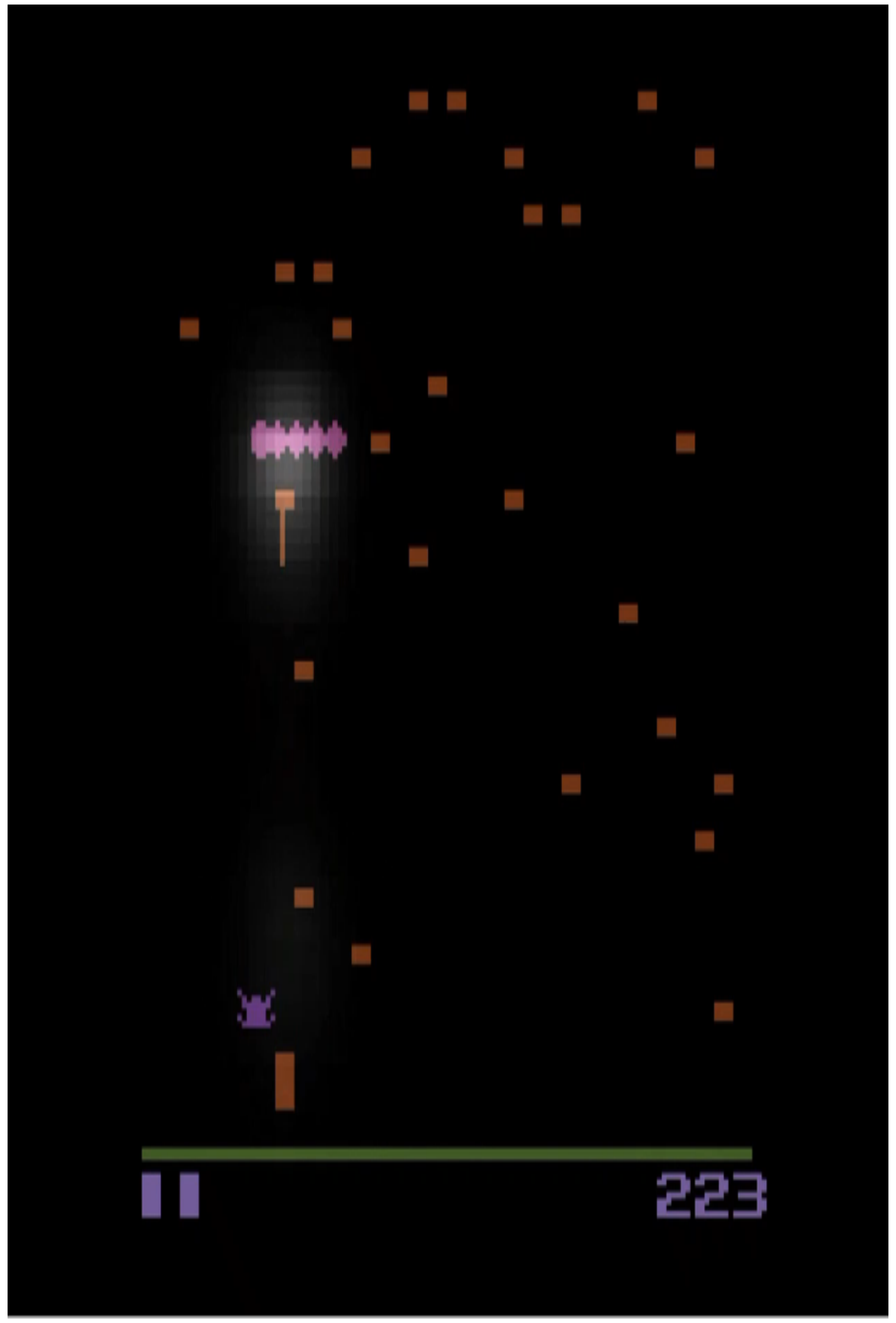
Visualizing learned Agent Policies

BC



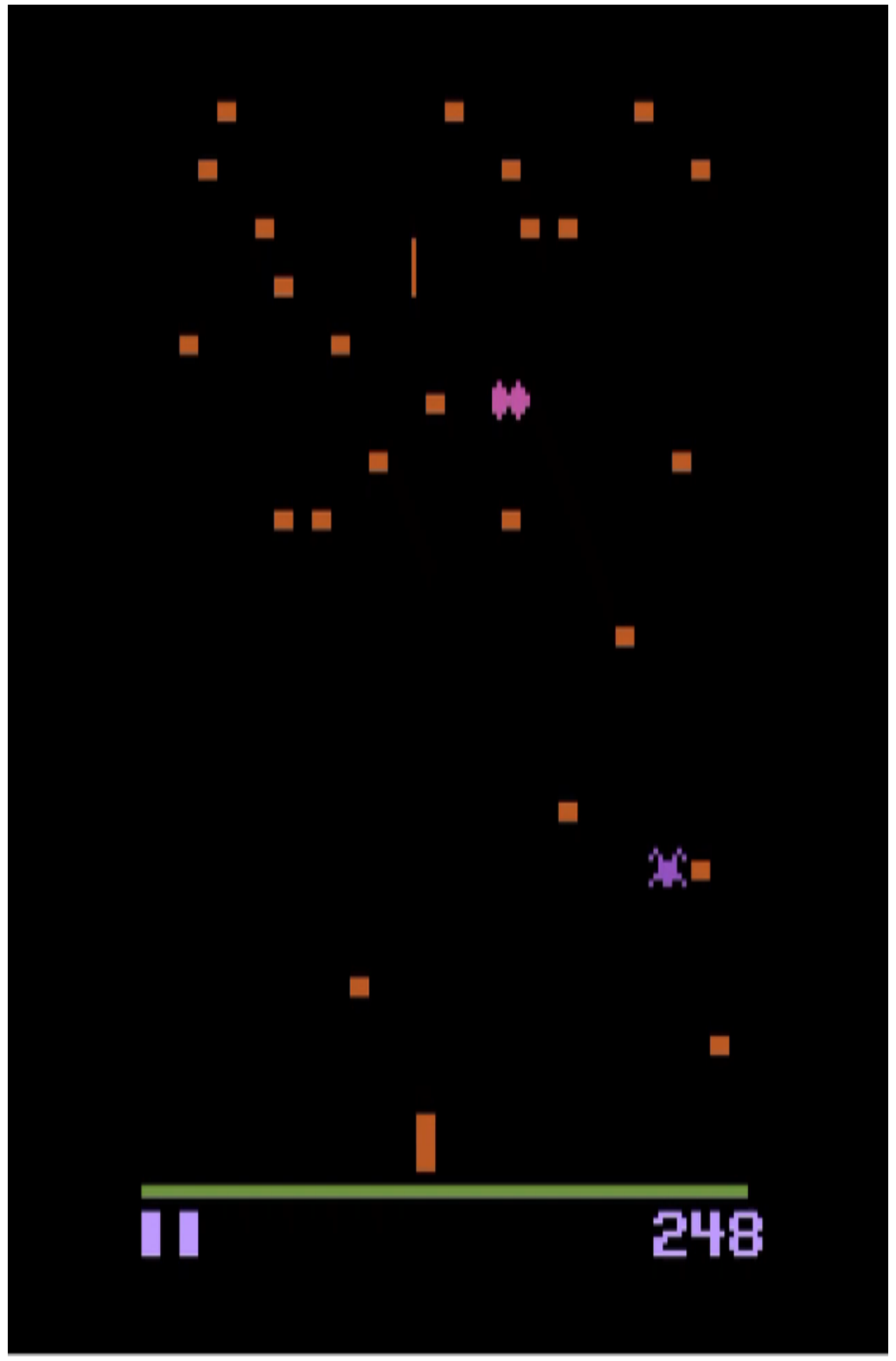
Does not learn to actively shoot the spider

AGIL



Shoots the spider when it comes directly above the agent

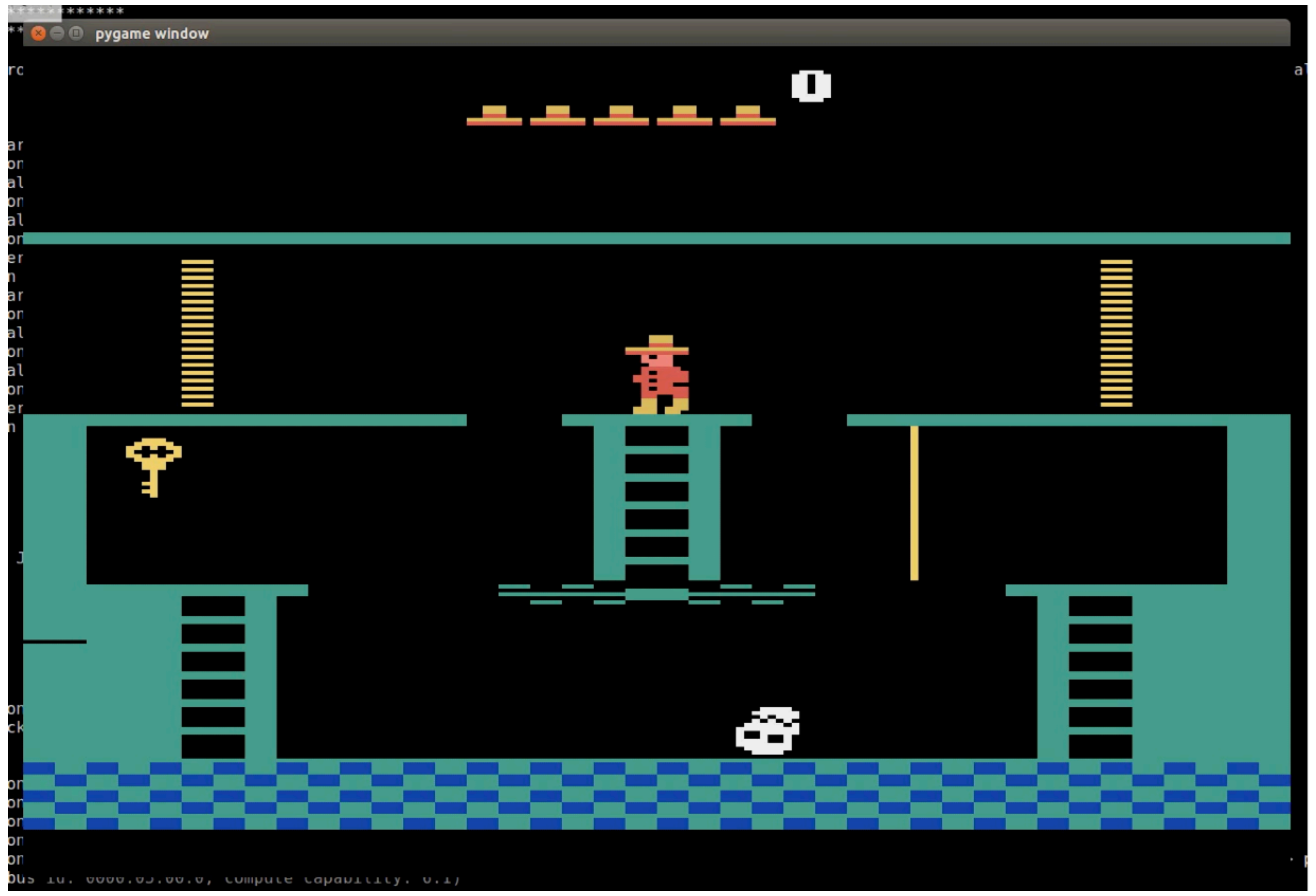
BC +CGL



Actively goes and shoots the spider

Visualizing CGL agent policies

BC

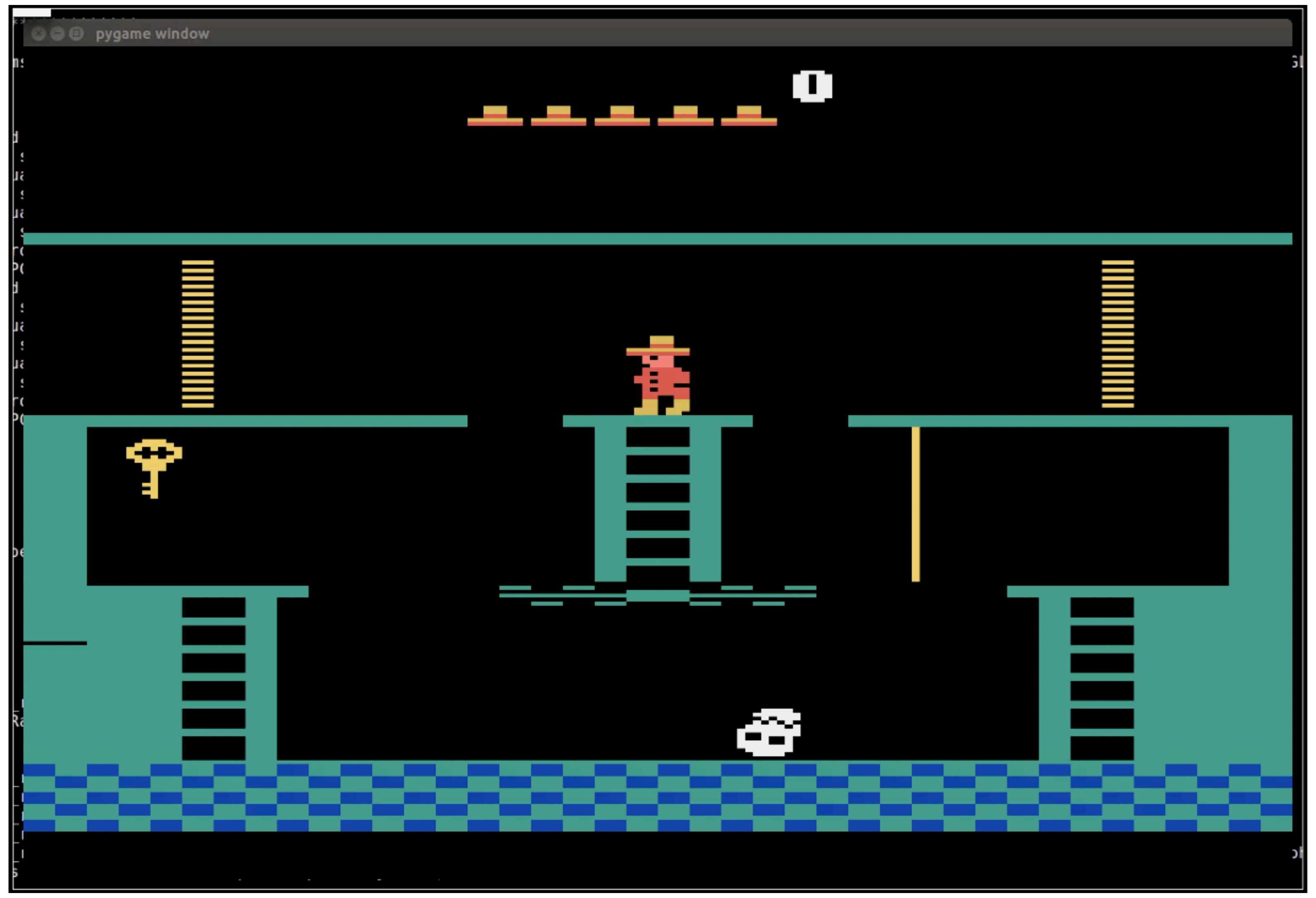


Unable to hop over the skull

AGIL

Unable to hop over the skull

BC +CGL



Learns to hop over the skull and advance ahead in the game

Understanding the Performance Gains of CGL

Can CGL reduce causal confusion for Imitation Learning methods?

Correct Causal Identification

P. de Haan, D. Jayaraman, & S. Levine. Causal confusion in imitation learning. NeurIPS 2019.

CGL reduces causal confusion compared to baseline BC algorithm

Confounded images with correlated past actions as part of the state space

(a) Breakout (b) Asterix (c) Demon Attack (d) Freeway

P. de Haan, D. Jayaraman, & S. Levine. Causal confusion in imitation learning. NeurIPS 2019.

CGL reduces causal confusion compared to baseline BC algorithm

CGL suffers less with confounded data and hence reduces causal confusion compared to BC

Algorithm tested with confounded images	Performance reduction with confounded images (lower is better)
BC [confounded] v/s BC [original]	-47.8 %
BC+CGL [confounded] v/s BC+CGL [original]	**-34.0 %**

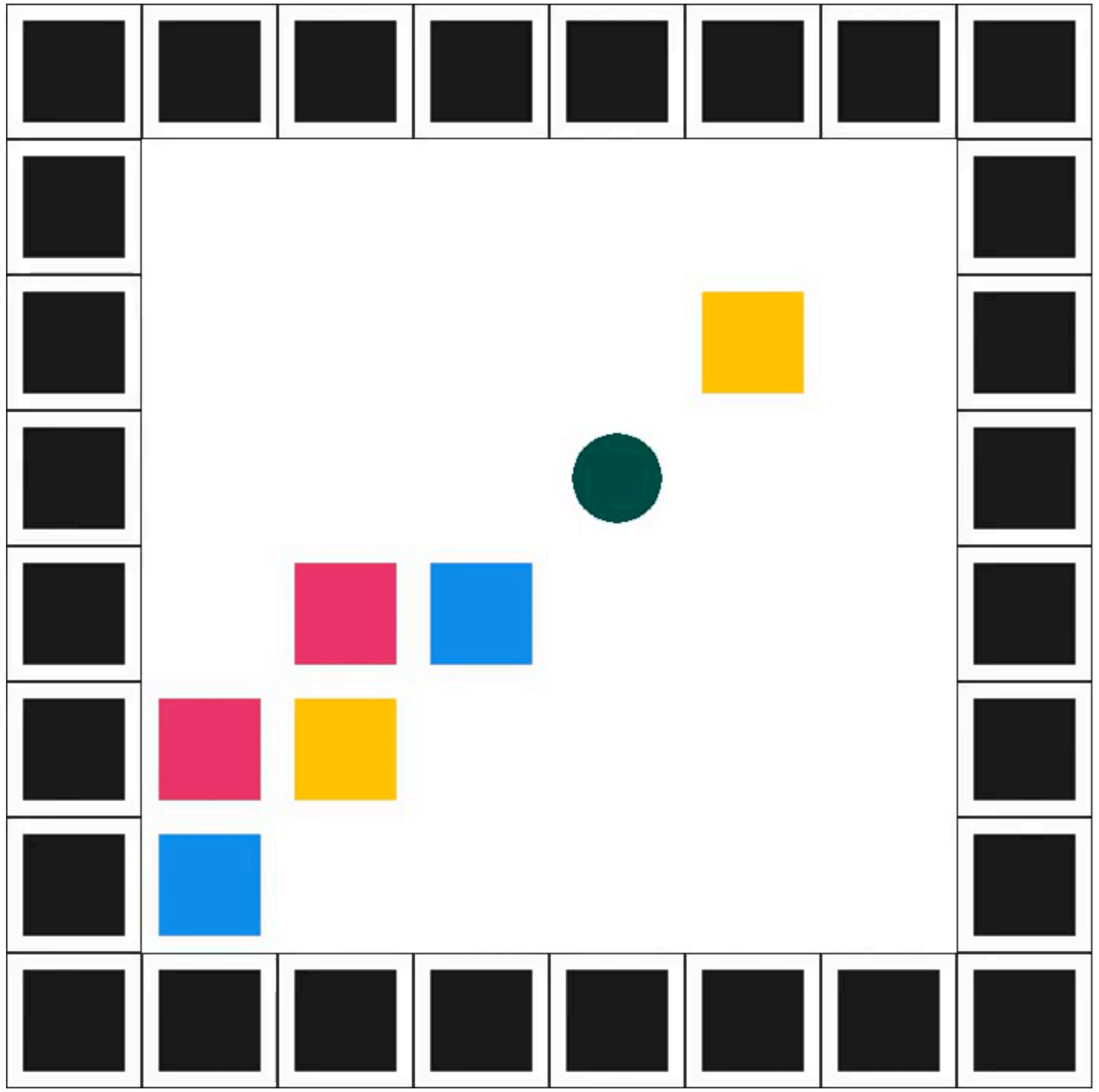
BC+CGL outperforms BC trained with confounded data by 571%

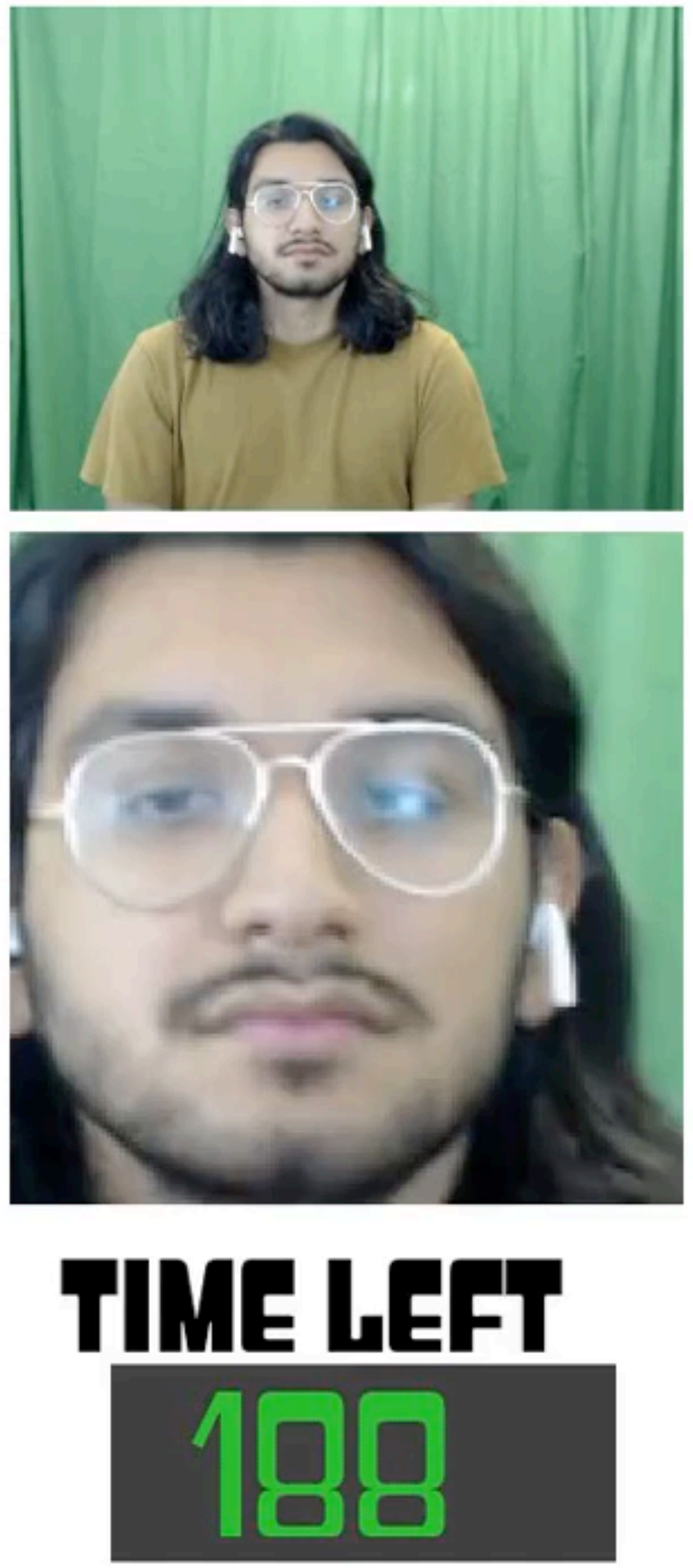
Saran, A., R. Zhang, E. S. Short, and S. Niekum2020. Efficiently guiding imitation learning algorithms with human gaze. AAMAS 2021.

Implicit human feedback: Facial Reactions

- Occurs naturally
- Is not necessarily intended to influence behavior
- Can be used with no additional burden on user

EMPATHIC: Learning from implicit feedback





Steps of EMPATHIC:

- Incentivize human participant
- Collected reaction data under known GT reward (or other task statistic of interest)
- Learn human reward model (or other task statistic)
- Transfer to new tasks

Y. Cui, Q. Zhang, A. Allievi, P. Stone, S. Niekum, and W. Knox.
[The EMPATHIC Framework for Task Learning from Implicit Human Feedback].
Conference on Robot Learning (CoRL), November 2020.

Task Domains

Robotaxi

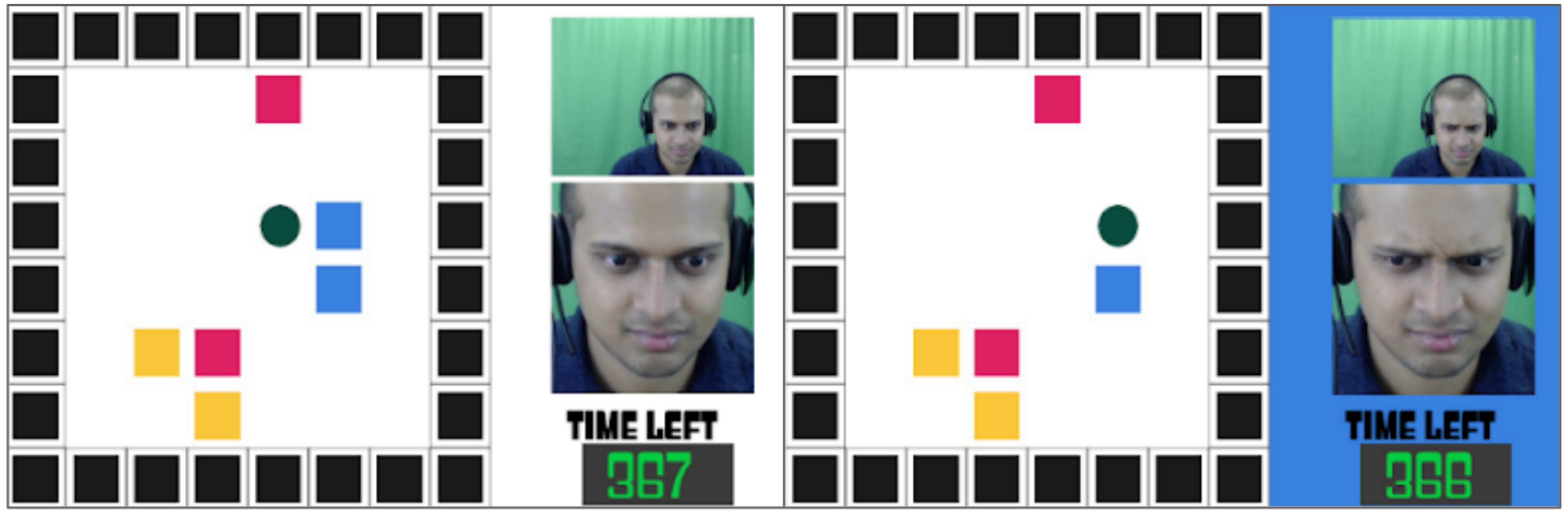
Robotic Trash Sorting

How hard is this problem?

Is there enough information to learn from implicit human feedback?

- Human proxy test
- Facial annotation data

	Avg. τ	p-value
Human Proxies	.569	.004
	.216	.185
	.098	.319
	-.176	.179
	.255	.123
	.294	.059
Avg.	.209	.078

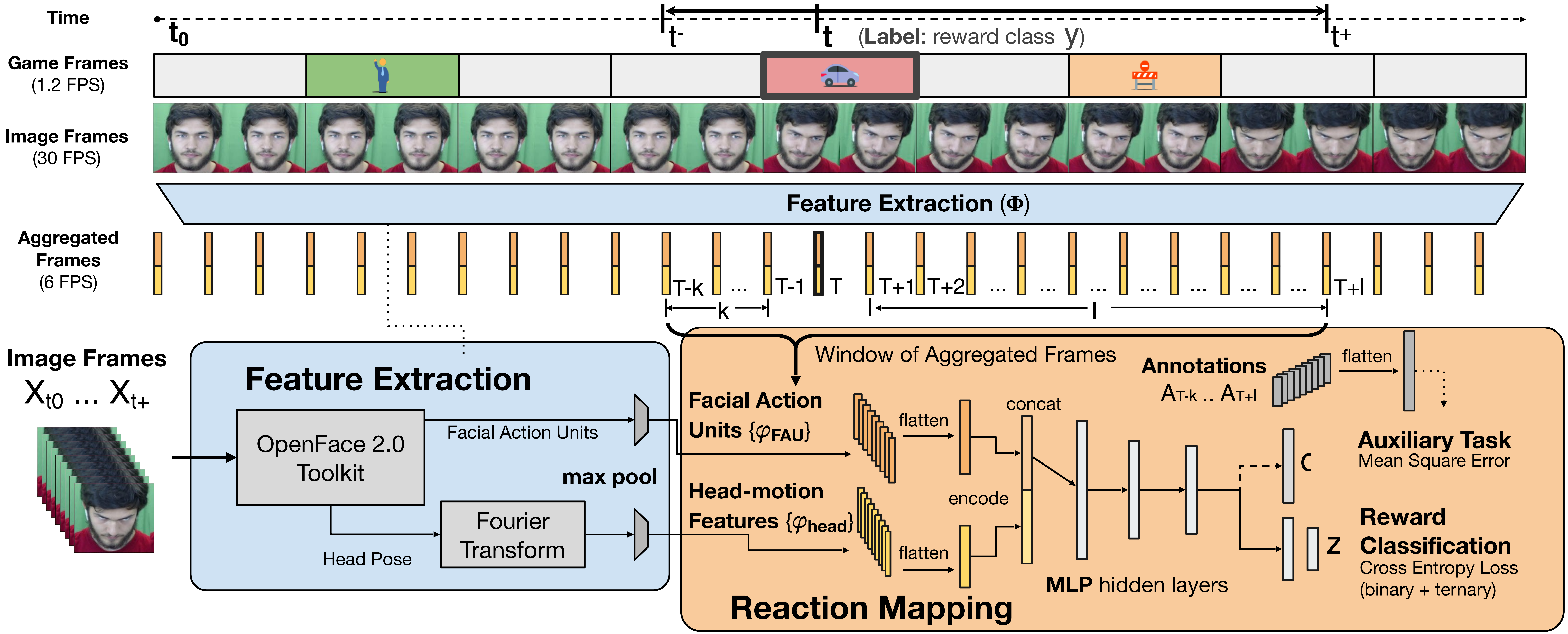


timestep **t** timestep **t+1**

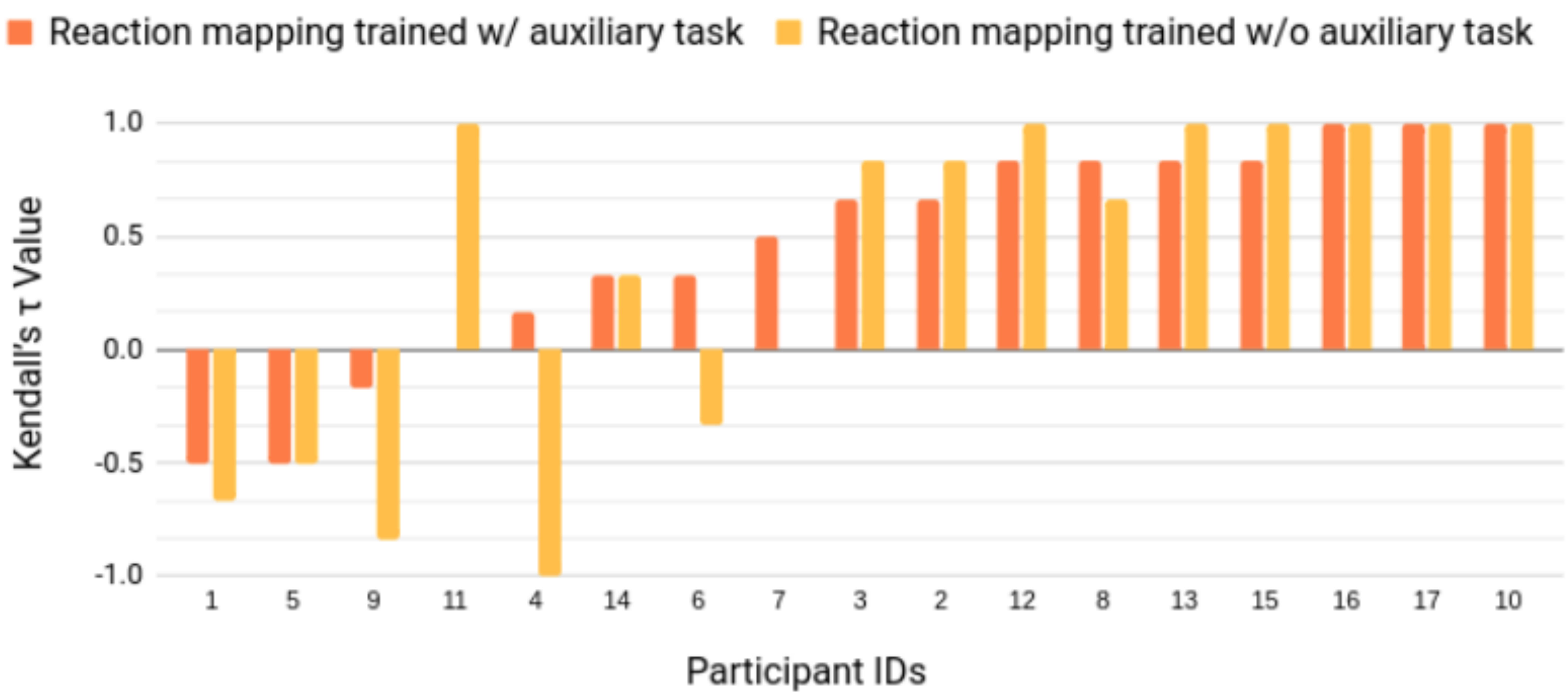

AnnotationTool
frame_number
head_nod
head_shake
eye_roll
smile
pout
eyebrow_raise
eyebrow_frown
923/5436
OFF
OFF
OFF
OFF
OFF
OFF
OFF
sentiment
positive
neutral
negative
Bulk Frame Annotation
Apply Annotation to Bulk Frames
start index: 601.0
+
-
end index: 923.0
+
-
play >>
stop ||
< prev frame
next frame >
Import Existing Annotations
Export Annotations

Analyzing Annotated Facial Gestures

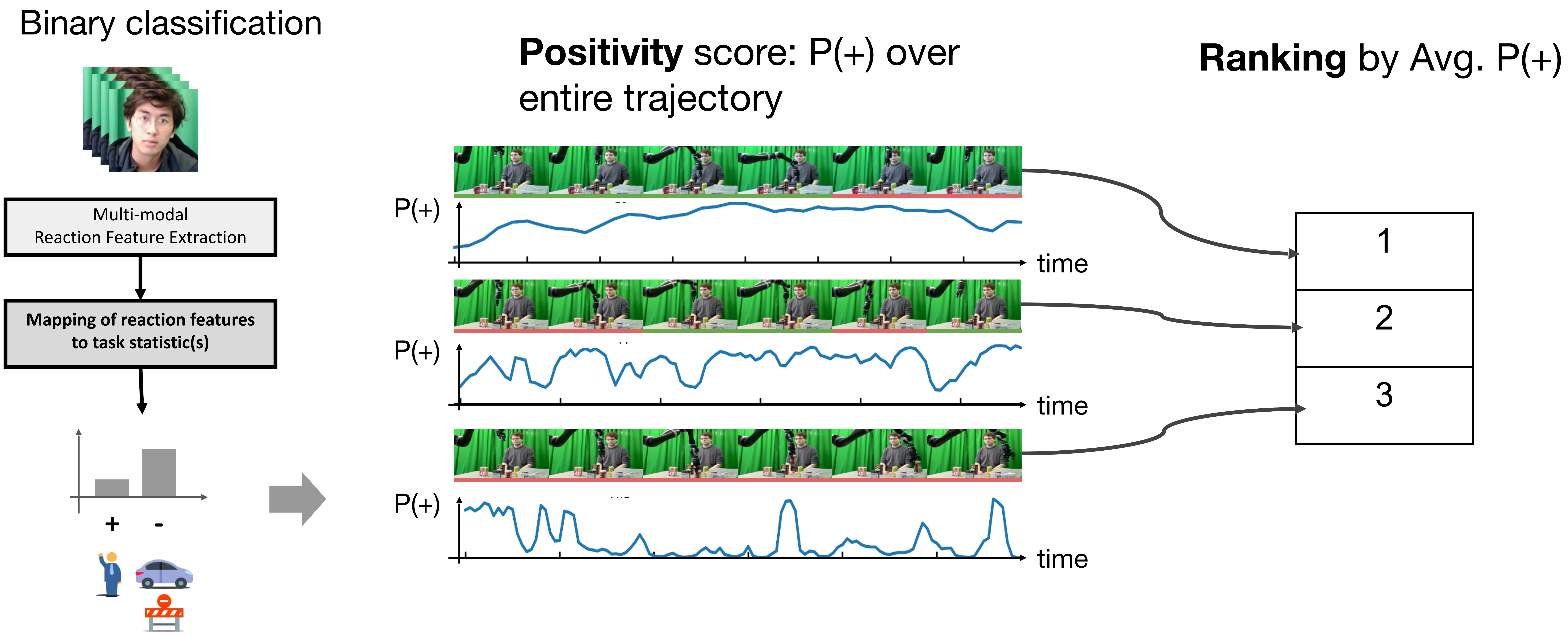
Learning the Reaction Mapping



Reward Ranking Prediction Performance

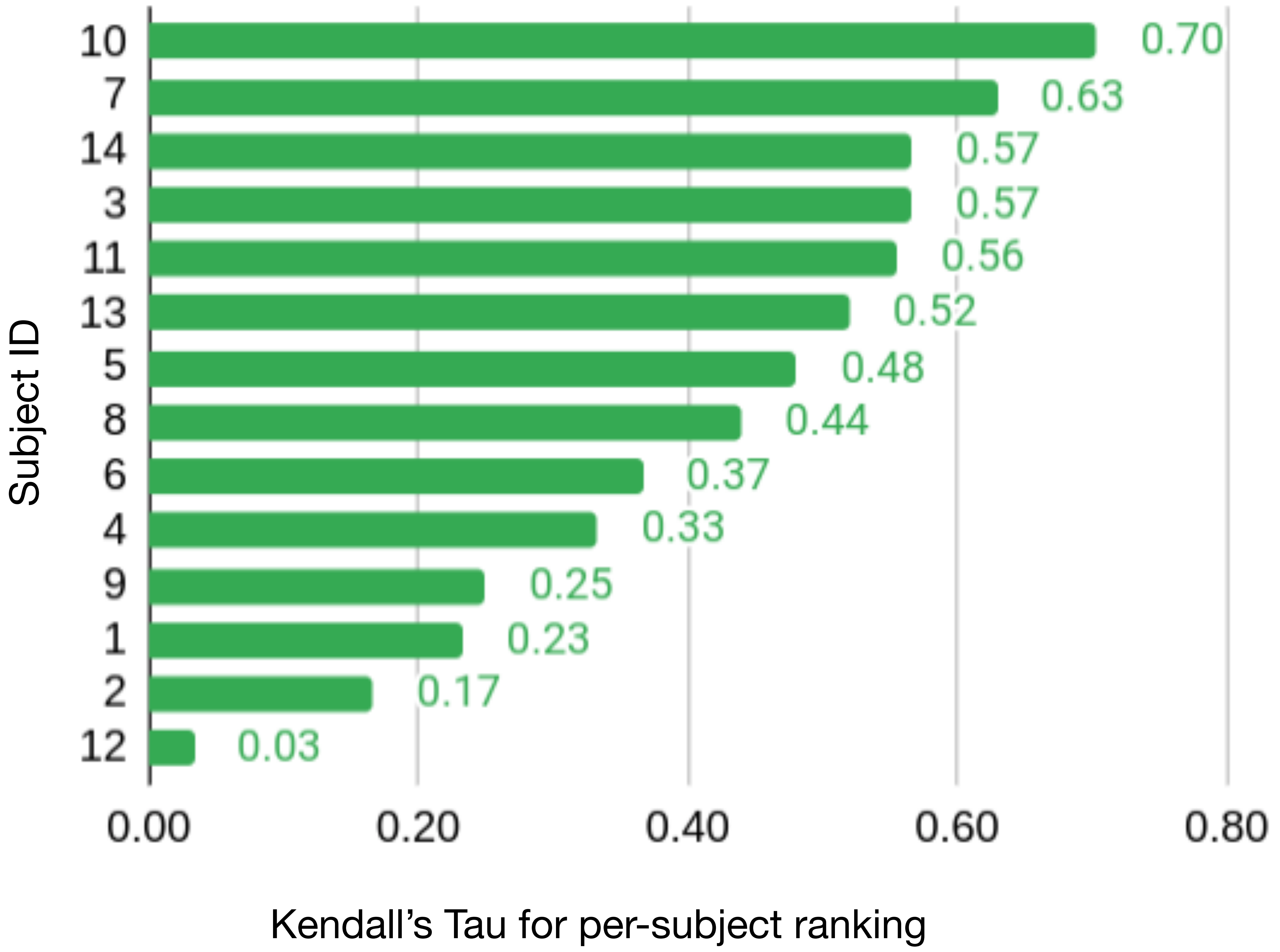


How to leverage the learned mapping from Robotaxi?



EMPATHIC: Learning from implicit feedback — deployment

Robotic Trash Sorting Performance



red bottle then can	0.163
waterloo can	0.156
white can	0.138
red bottle	0.107
brown box	0.097
yellow can	0.087
green box	0.073
white box	0.045

Overall Ranking (avg. positivity)