OHBM 2026

A naturalistic reinforcement learning paradigm for characterizing real-world risky behavior

Jinwoo Jeong[1,2] and Woo-Young Ahn[1,3,4]

[1]Department of Psychology, Seoul National University; [2]School of Computer Science, Georgia Institute of Technology; [3]Department of Brain and Cognitive Sciences, Seoul National University; [4]AI Institute, Seoul National University

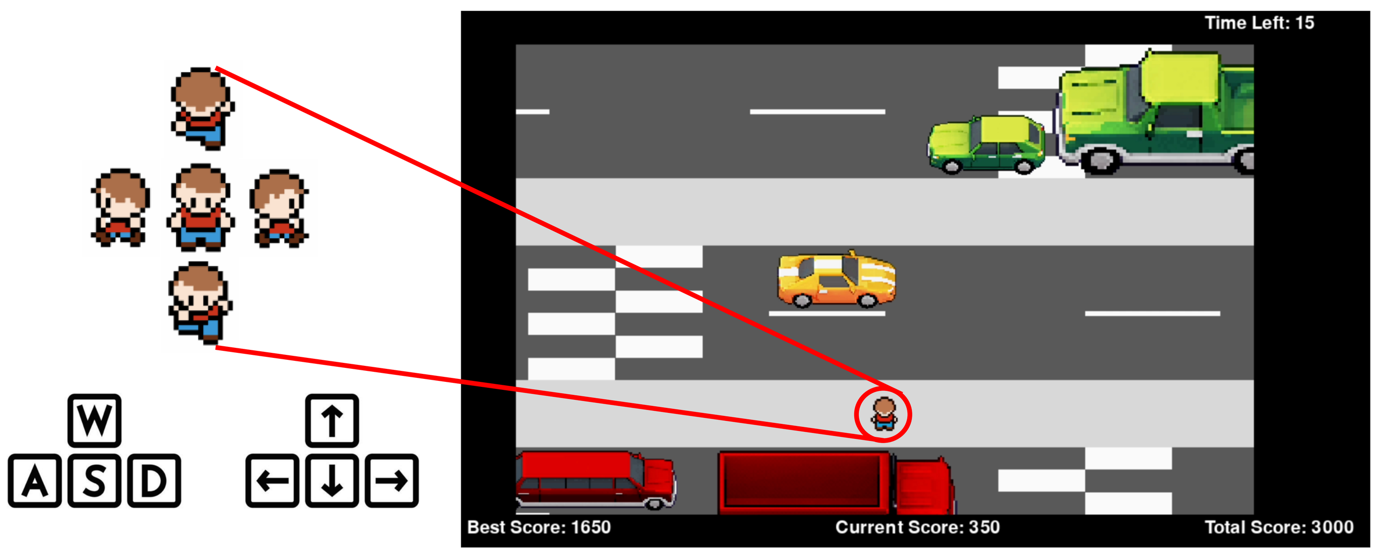
Professional Profile

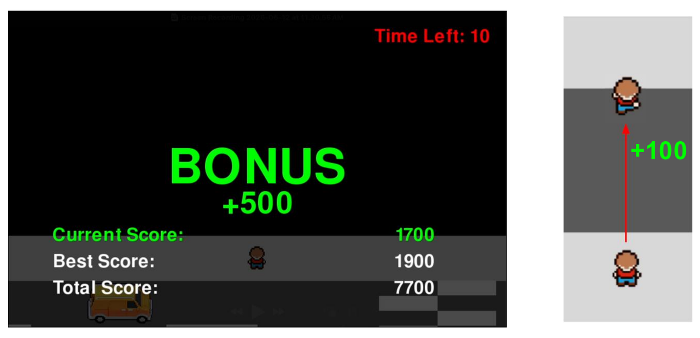
Introduction

- **Naturalistic reinforcement learning** suggests that real-world behavior emerges from continuous, dynamic, and context-dependent interactions (Wise et al., 2024).
- Recent studies have shown that impulsive behaviors observed during a **real-time driving task** can reflect individual differences in trait-level impulsivity and are linked to reward-related neural responses (Lee et al., 2024; Lee et al., 2026).
- **Road-crossing behavior** has also been used as a naturalistic paradigm to assess risk-taking propensity, but prior work has not fully captured the continuous dynamics of real-time decision making (Herrero-Fernández et al., 2016; Rubio et al., 2010).
- To address this gap, we developed the **crosswalk task**, a naturalistic real-time decision-making task designed to capture individual differences in risk-taking.

Crosswalk Task

- **Goal**: cross all the roads and reach the end of the map as fast as possible

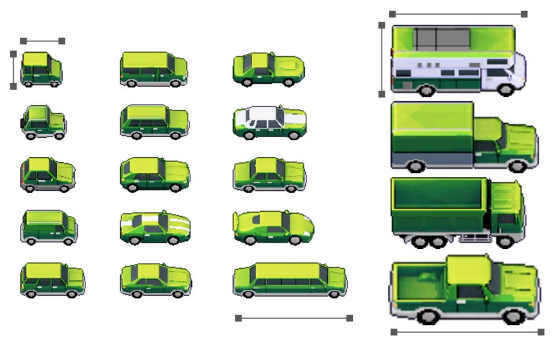




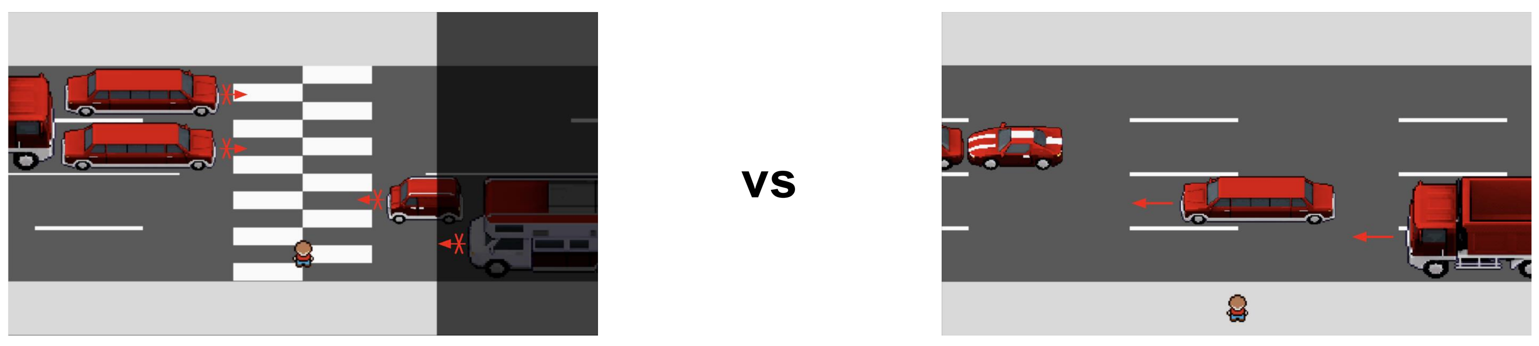
- Road-crossing reward
- Episode completion bonus

- Collision penalty based on car color
- One car color per road

- Episode time limit: 30 seconds



- Size variation => varying crossing difficulty



Crosswalk Activation: risk-averse strategy

Jaywalking: risk-taking strategy

Methods

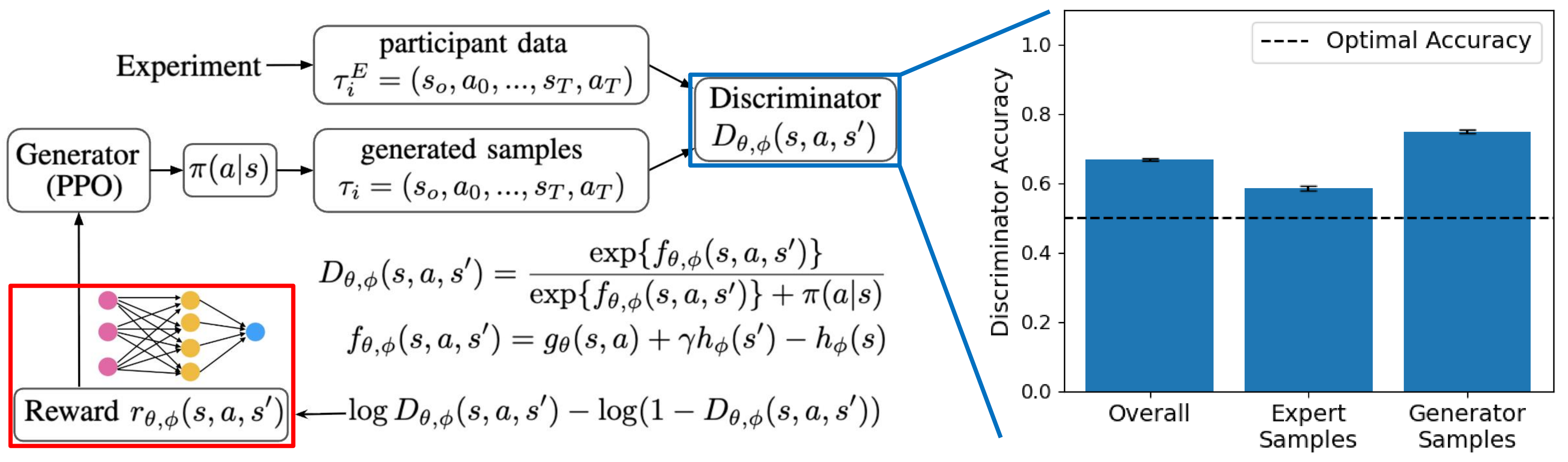
Data Collection

	Study 1 (N = 55)	Study 2 (N = 254)
Participants	Korean adults	U.S. adults
Sex	23 males / 32 females	107 males / 140 females
Age (M±SD)	25.2 ± 5.2	37.4 ± 11.0
Experiment Setting	In-person (SNU lab)	Online (Prolific)
Collection Period	2025.10.03 – 2025.12.05	2026.03.23 – 2026.04.07

Adversarial Inverse Reinforcement Learning

- One AIRL model trained per participant (N = 309)
- **Reward networks** used to compute state–action rewards ("IRL rewards")

⇒ **IRL rewards** used to identify context-specific indicators of risk-taking



Elastic Net Regression

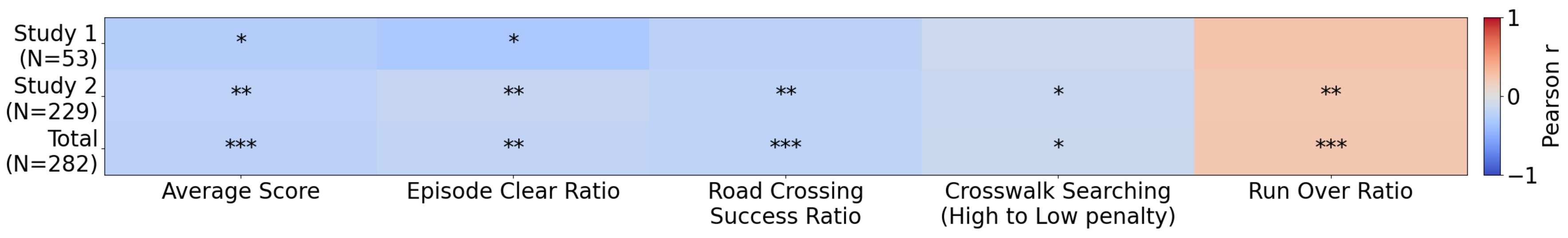
- Goal: **predict DOSPERT*** total scores using **crosswalk task features**
- Model: Elastic Net (α = 0.5) for **feature selection** and **prediction**
- Evaluation: 80/20 train–test split + 5-fold cross-validation

**DOSPERT: 30-item self-reported risk-taking measure*

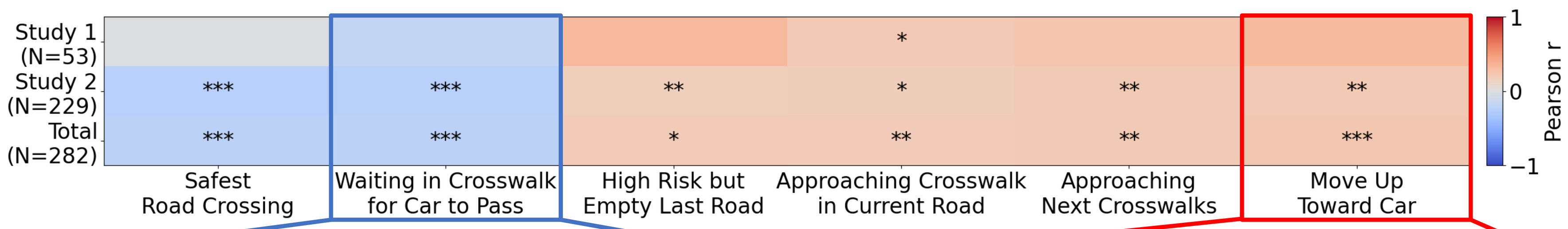
Results

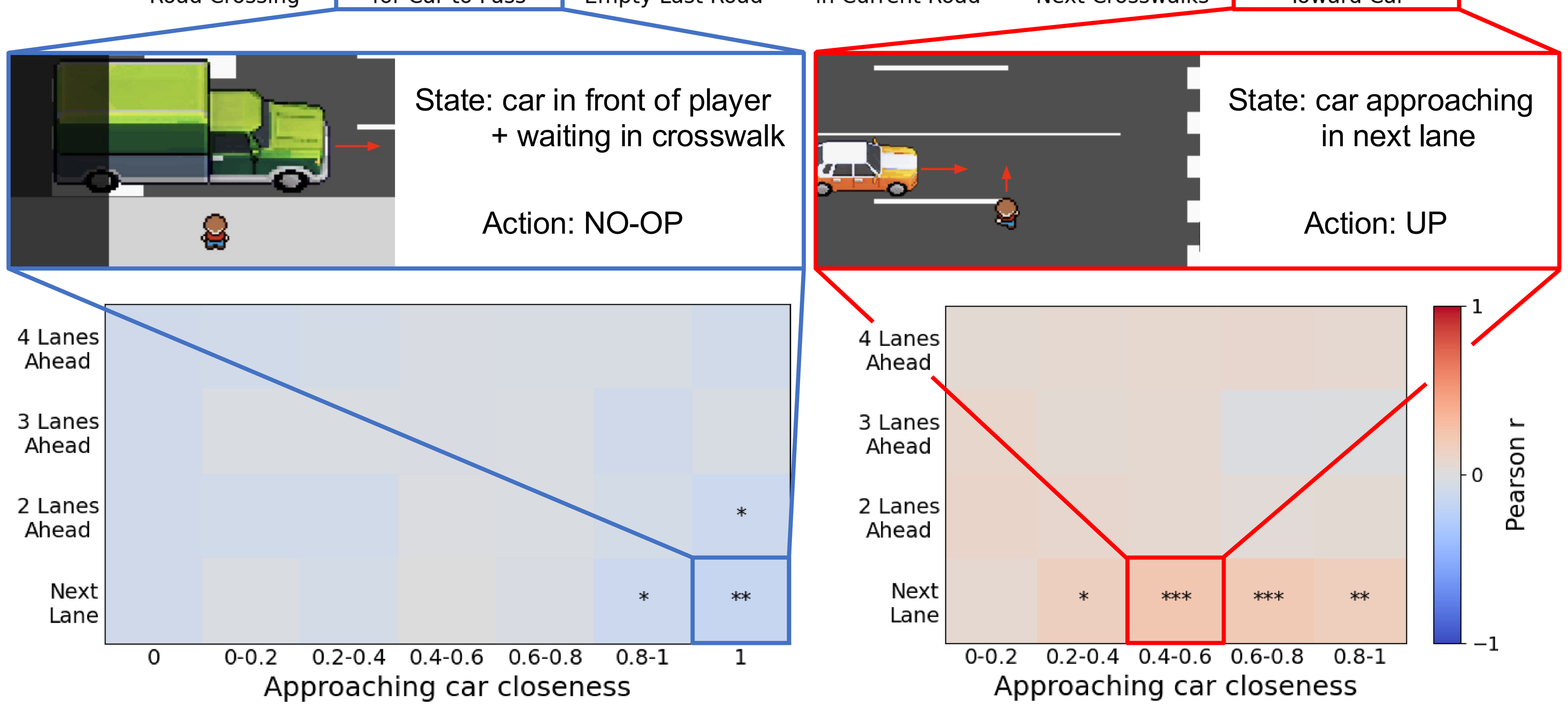
Identifying Indicators of Risk-Taking

- Goal: Identify crosswalk task features correlated with DOSPERT
- **Behavioral features**: task-level summary statistics
  - e.g., run-over ratio = car collisions / total episodes;
    road-crossing success ratio = successful crossings / encountered roads



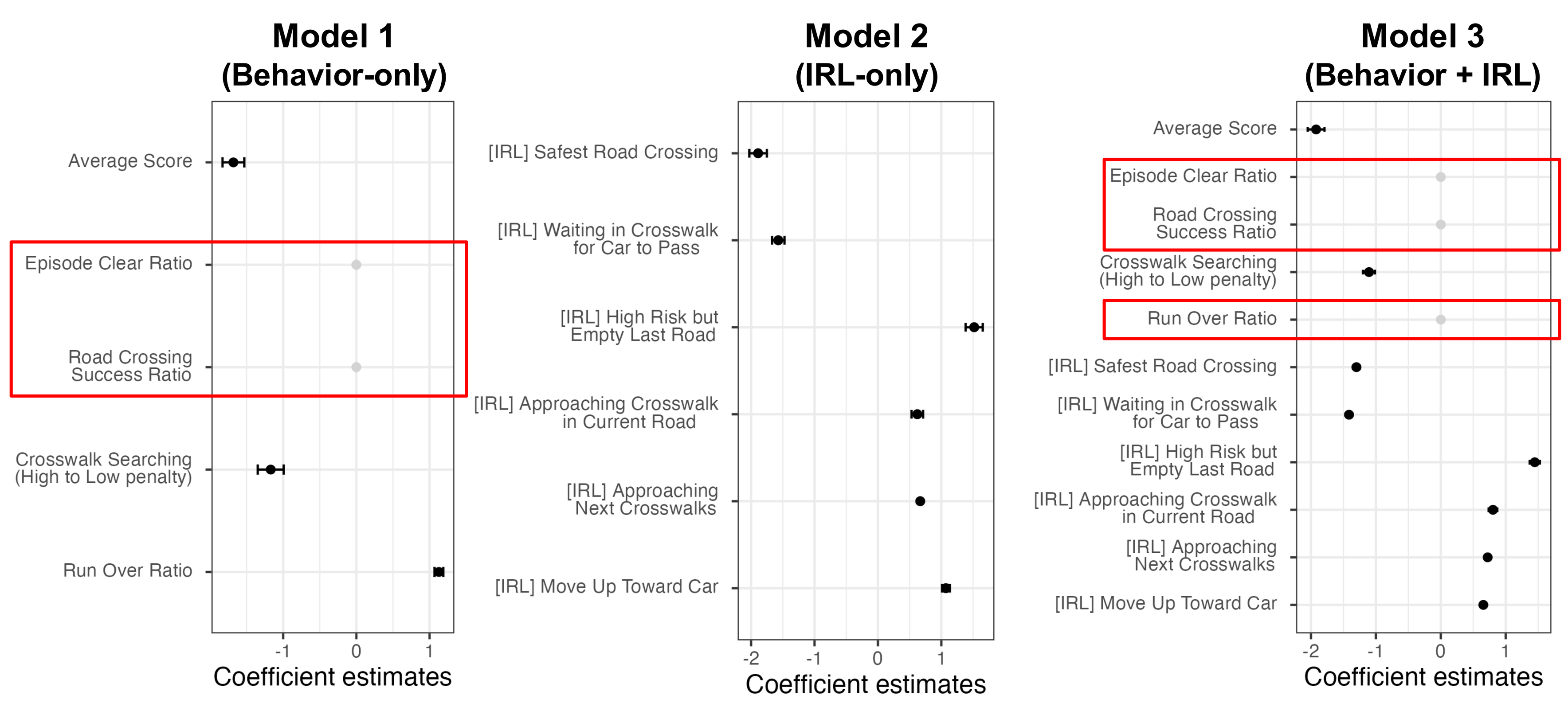
- **IRL features**: IRL reward estimates for specific state-action conditions





**Car closeness: 0 = no visible car in the lane; 1 = car in front of the player*

Elastic Net Model Comparison



	Model 1	Model 2	Model 3
Feature Selection	Behavioral: 3/5	**IRL: 6/6**	**IRL: 6/6** Behavioral: 2/5
Pearson r (Train)	0.2632	0.3640	0.4153
Pearson r (Test)	0.2237	0.3067	0.3345

- IRL features were **more likely to be selected** than behavioral features.
- Models with IRL features showed **better predictive performance** than Model 1

=> **IRL features** provided more consistent and informative predictors of risk-taking

Conclusion

- A **real-time task** designed to capture both risk-taking and risk-averse behaviors can yield **features** associated with self-reported risk propensity.
- **IRL-derived context-specific indicators** can better capture individual differences in risk-taking behavior than aggregate task-level summaries.
- These findings highlight the crosswalk task's potential as a meaningful behavioral paradigm for studying real-time risk-taking in future neuroimaging studies.

References

Fu, J., Luo, K., & Levine, S. (2018). Learning robust rewards with adversarial inverse reinforcement learning. *In Proceedings of the International Conference on Learning Representations*.

Herrero-Fernández, D., Macía-Guerrero, P., Silvano-Chaparro, L., Merino, L., & Jenchura, E. C. (2016). Risky behavior in young adult pedestrians: Personality determinants, correlates with risk perception, and gender differences. *Transportation research part F: traffic psychology and behaviour, 36*, 14-24.

Lee, S. H., Song, M. S., Oh, M., & Ahn, W.-Y. (2024). Bridging the gap between self-report and behavioral laboratory measures: A real-time driving task with inverse reinforcement learning. *Psychological Science, 35*(4), 345–357.

Lee, S. H., Chung, C., Oh, M., & Ahn, W.-Y. (2026) Inverse reinforcement learning reveals action-oriented value signals in naturalistic decision making. *bioRxiv*.

Rubio, V. J., Hernández, J. M., Zaldívar, F., Márquez, O., & Santacreu, J. (2010). Can We Predict Risk-Taking Behavior?. *European Journal of Psychological Assessment.*

Wise, T., Emery, K., & Radulescu, A. (2024). Naturalistic reinforcement learning. *Trends in Cognitive Sciences, 28*(2), 144-158.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 67*(2), 301-320.

- Arrived at France to present a poster at #OHBM2026 (06.14 - 06.18).

- Poster 2078: A naturalistic reinforcement learning paradigm for characterizing real-world risky behavior

- developed a custom Gymnasium environment and real-time decision-making task for cognitive research.

- collected and analyzed behavioral data from 300+ participants in an online experiment using Prolific.

- trained inverse reinforcement learning models to derive context-specific indicators of risk-taking.

Arrived at France to atte
Braining Mapping.

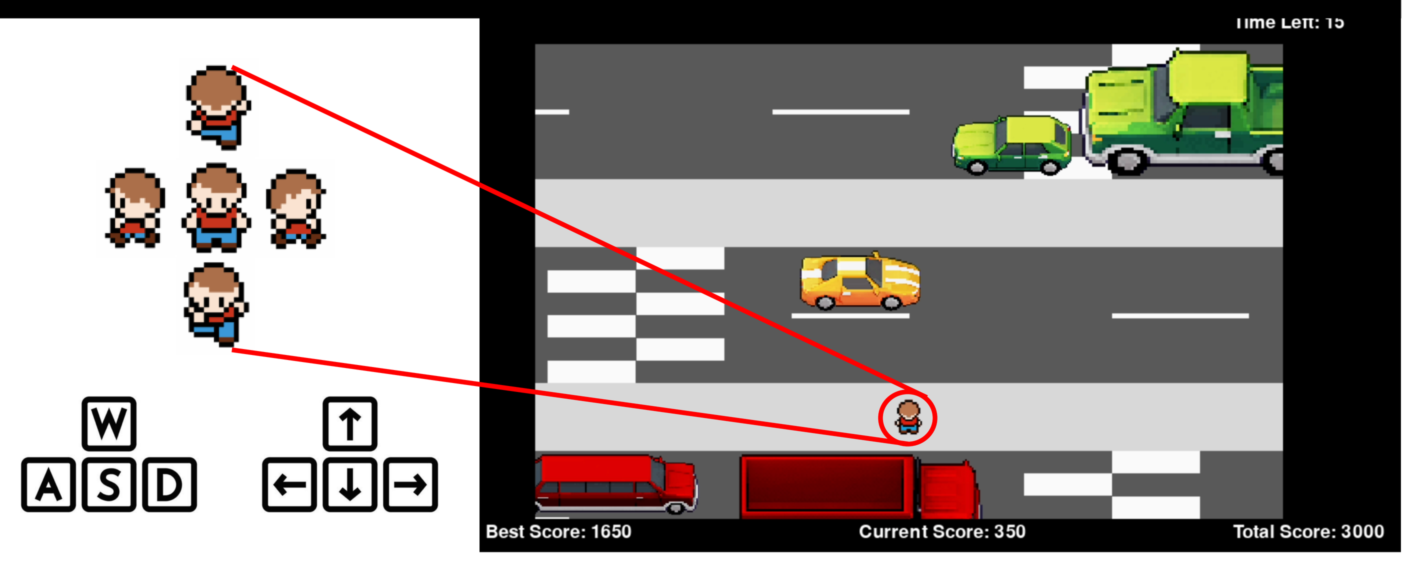
This year, I will be prese

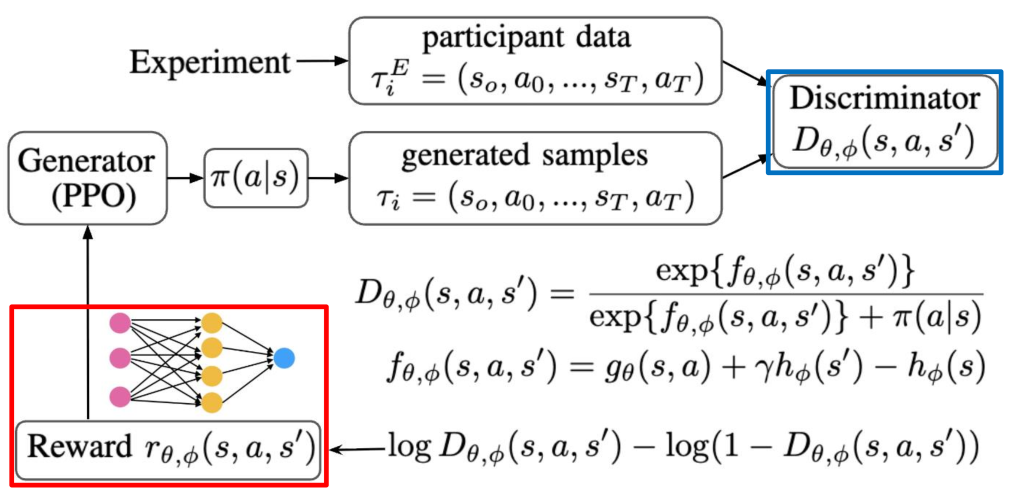
A naturalistic reinforcem
characterizing real-world

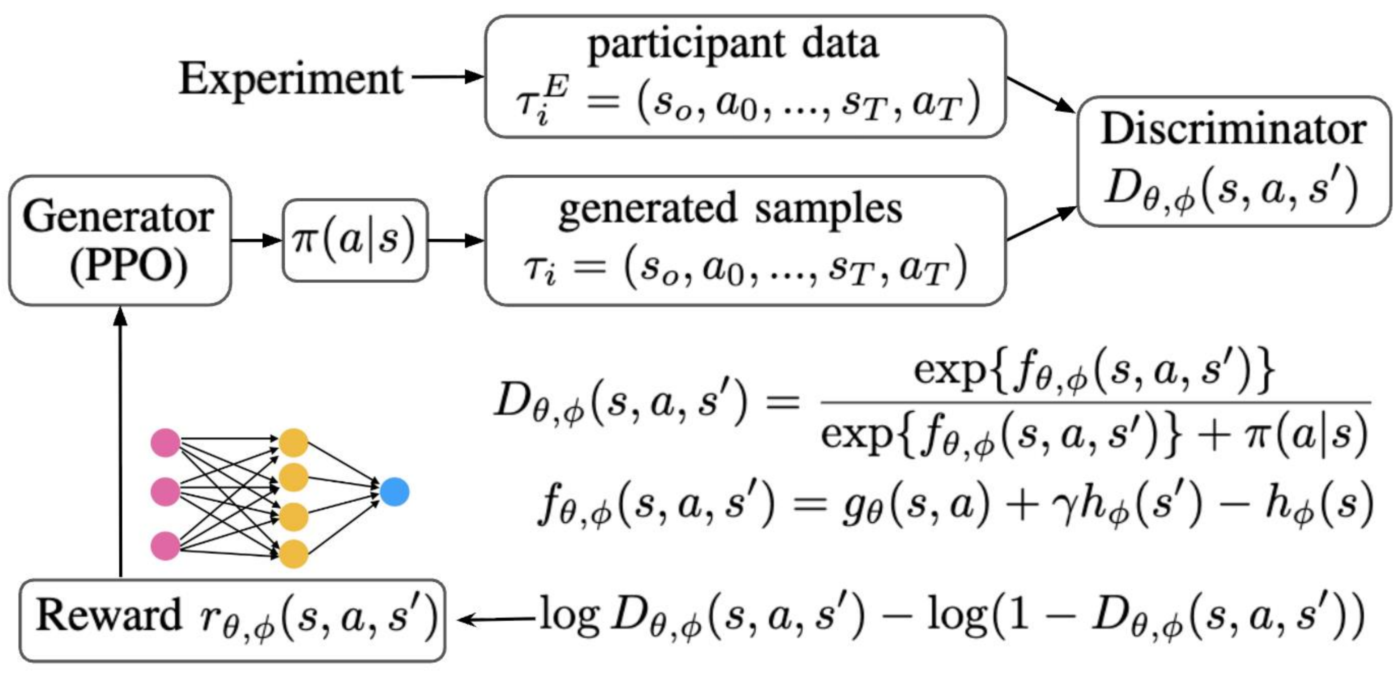
Applying deep inverse reinforcement learning to analyze individual differences in human behavior

Arrived at France to present a poster at OHBM 2026 (06.14 - 06.18).
Arrived at France to present a poster at OHBM 2026 (06.14 - 06.18).
Arrived at Bordeaux, France to present a poster at OHBM 2026 (06.14 - 06.18).

Poster number = 2078
First Stand-by Time = Wednesday, June 17 | 13:45-14:45
Second Stand-by Time = Thursday, June 18 | 14:45-15:45
Title = A naturalistic reinforcement learning paradigm for characterizing real-world risky behavior
Primary Category = Modeling and Analysis Methods







- Arrived at France to present a poster at OHBM 2026 (06.14 - 06.18).

- Project: A naturalistic reinforcement learning paradigm for characterizing real-world risky behavior

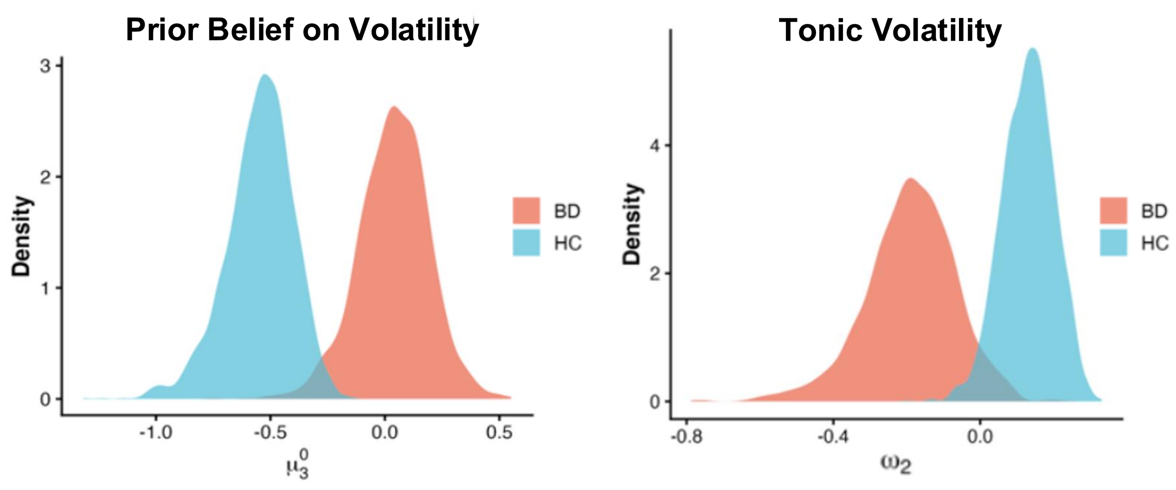
- developed a custom Gymnasium environment and real-time decision-making task for cognitive research.

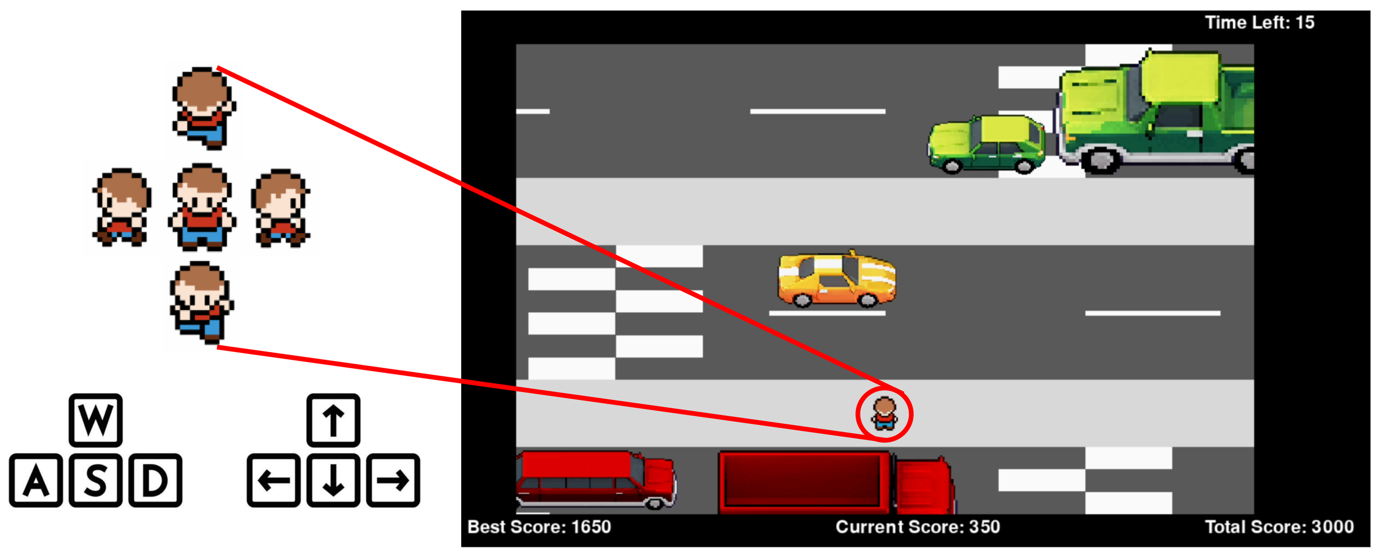
- collected and analyzed behavioral data from 300+ participants in an online experiment using Prolific.

- trained inverse reinforcement learning models to derive context-specific indicators of risk-taking.

Previously gave an aral presentation at Korean Society for Digital Therapeutics 2026 Spring Conference

We will start with publications





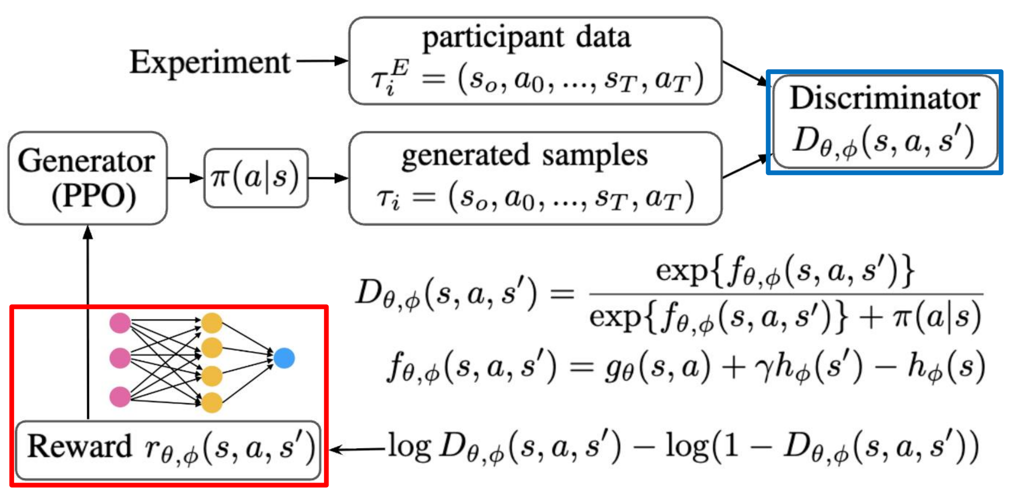
Arrived at France to attend Organization of Human Braining Mapping.

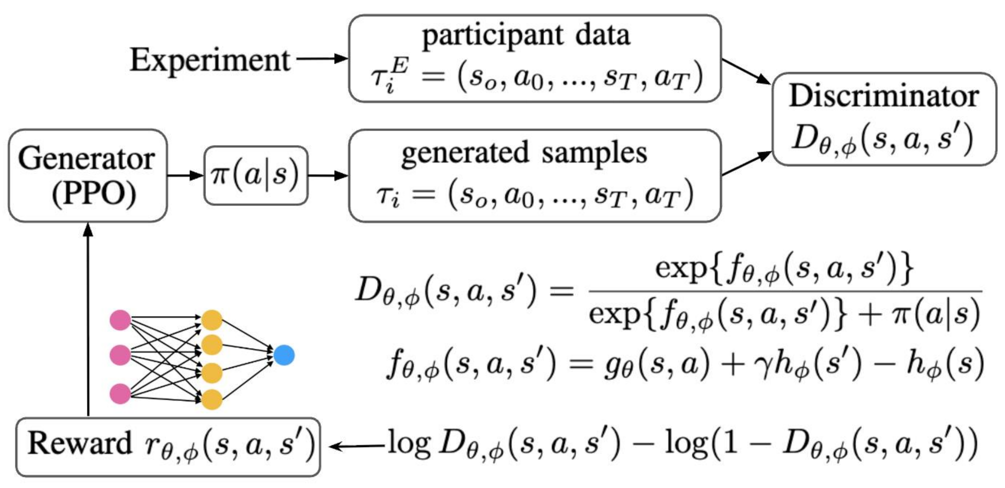
This year, I will be presenting a poster on

A naturalistic reinforcement learning paradigm for characterizing real-world risky behavior

Applying deep inverse reinforcement learning to analyze individual differences in human behavior

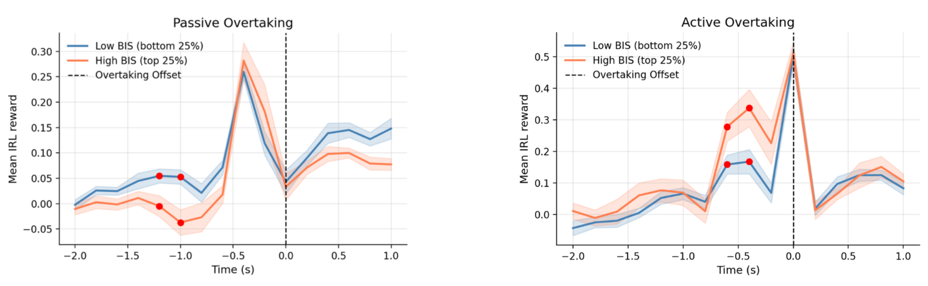
Arrived at France to present a poster at OHBM 2026 (06.14 - 06.18).
Arrived at France to present a poster at OHBM 2026 (06.14 - 06.18).
Arrived at Bordeaux, France to present a poster at OHBM 2026 (06.14 - 06.18).





IRL 결과 1: 충동성이 높을수록 위험한 보상에 더 높은 가치 부여

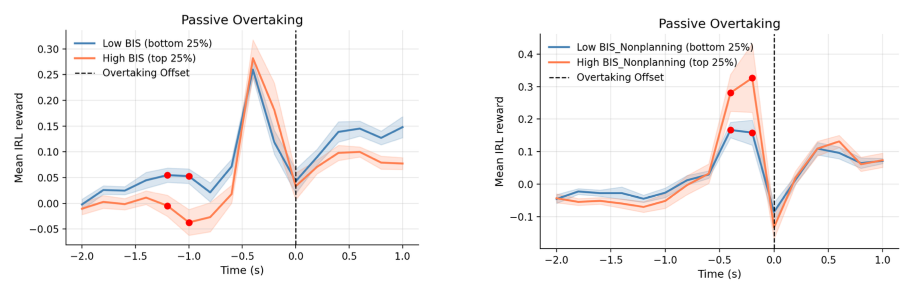
정적 환경 (static environment)



정적 환경에서 **수동적 추월** 이전 **충동성이 낮을수록 높은 보상**을 보이지만 (r = -0.36; BF = 4.3),
능동적 추월 이전 **충동성이 높을수록 높은 보상**을 보임 (r = 0.34; BF = 3.24).

A. 정적 환경 (static environment)

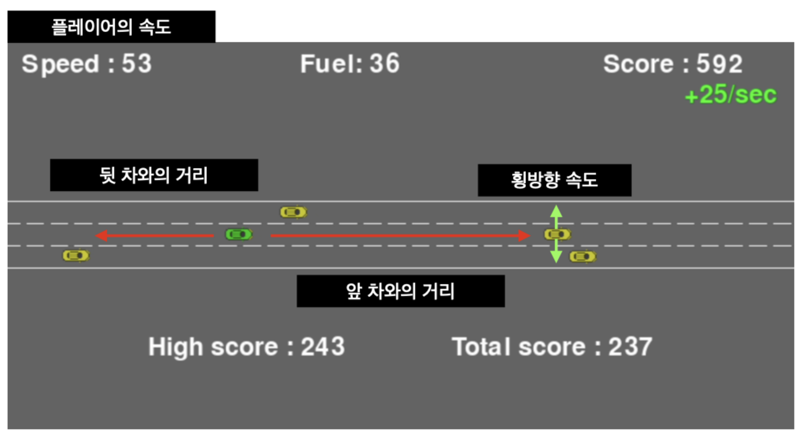
B. 동적 환경 (dynamic environment)

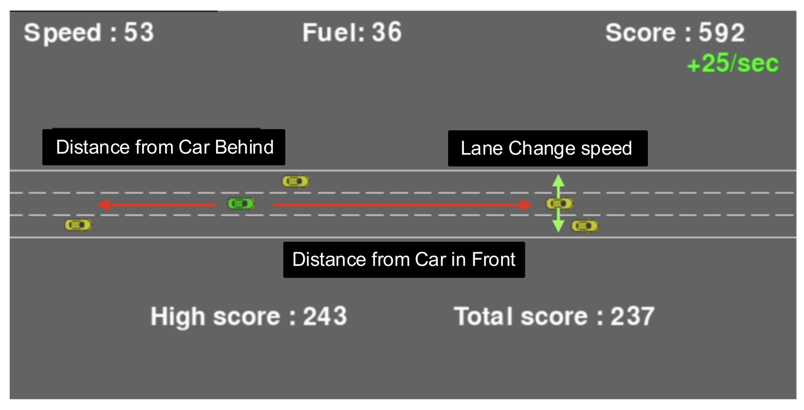


[정적 환경] 수동적 추월 시 **충동성이 낮을수록 높은 보상** (r = -0.36; BF = 4.3)

- 특히 운동 충동성과 연관 (t = -1.2 to -1.0s; r = -0.38; BF = 6.4)

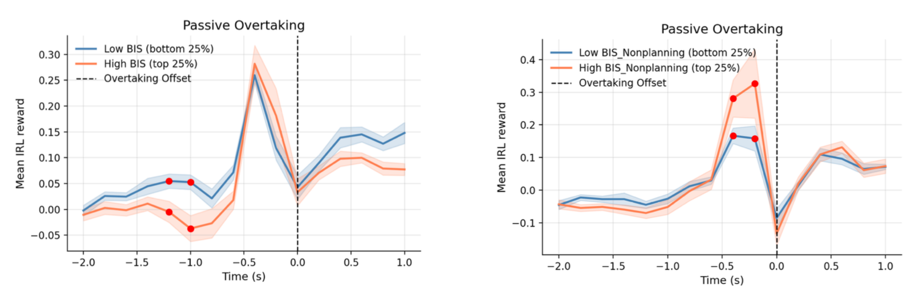
[동적 환경] **비계획 충동성이 높을수록 높은 보상** (r = -0.35; BF = 3.13).





A. 정적 환경 (static environment)

B. 동적 환경 (dynamic environment)



[정적 환경] 수동적 추월 시 **충동성이 낮을수록 높은 보상** (r = -0.36; BF = 4.3)

- 특히 운동 충동성과 연관 (t = -1.2 to -1.0s; r = -0.38; BF = 6.4)

[동적 환경] **비계획 충동성이 높을수록 높은 보상** (r = -0.35; BF = 3.13).