

Multi-agent reinforcement learning investigation of the mechanistic role of empathy on pro-environmental behavior

Jérôme Iste^{1,2,3,*}, Evann Rabeau^{1,2,*}, Ismael T. Freire³, Mehdi Khamassi³

¹Université Paris Cité, France; ² Sorbonne Université, France; ³ ISIR – Institut des Systèmes Intelligents et de Robotique

*Authors with equal participation



SORBONNE
UNIVERSITÉ



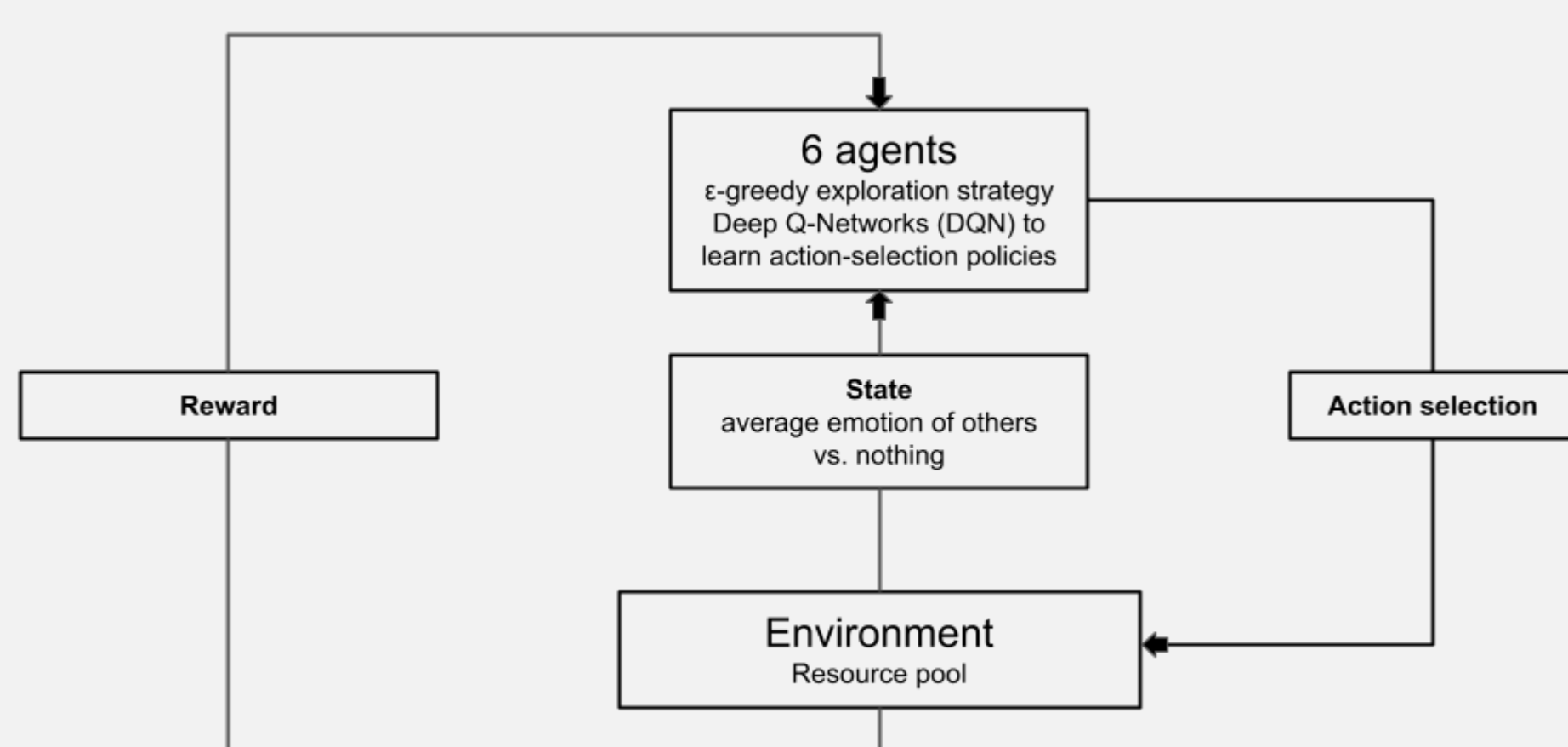
Introduction

- How can we avoid the Tragedy of the Commons (Hardin, 1968)?
 - Prior reinforcement learning simulation studies showed that modifications of agents' reward structures or observations can induce cooperation, conventions, and sustainability (Perolat et al., 2017; Wang et al., 2019; Zhu & Kirley, 2019).
- OBJECTIVE:**
- Using multi-agent reinforcement learning, we study the influence of the components of empathy (cognitive, affective, motivational) on common-pool resource dilemmas. We expect the conjunction of the ability to observe and be motivated by others' emotions to reduce overconsumption and increase fairness (Perolat et al., 2017; Wang et al., 2019; Zhu & Kirley, 2019).

Methods

ENVIRONMENT

- Renewable resource pool, that depends on the remaining stock
- Episode ends at depletion or 1000 steps (reset of resource stock)



EMPATHY IN THE RL FRAMEWORK

- Cognitive component:** agent's ability to observe a numerical value (the average emotions of others), which depends on their consumption
- Motivational component:** the importance (proportion) of others' rewards in the reward function of the agent

$$R_i^{\text{empathic}} = (1-\alpha) \cdot R_i^{\text{self}} + \alpha \cdot \frac{1}{N-1} \sum_{j \neq i} e_j$$

- R_i^{empathic} : reward received by the agent
- R_i^{self} : reward based on the personal history of resource access
- e_j : Emotion of agent j , computed via a linear mapping of other agents' histories
- $\alpha \in [0, 1]$: Weight of other agents' rewards (empathy factor)
- N : Total number of agents

EXPERIMENTAL CONDITIONS

STUDY 1: EMPATHY COMPONENT MATRIX

- 2 (level of empathy: 0 vs. 0.5) x 2 (see emotion: True vs. False)
- 22 simulations of 500 episodes each.

STUDY 2 : MULTIPLE ALPHA

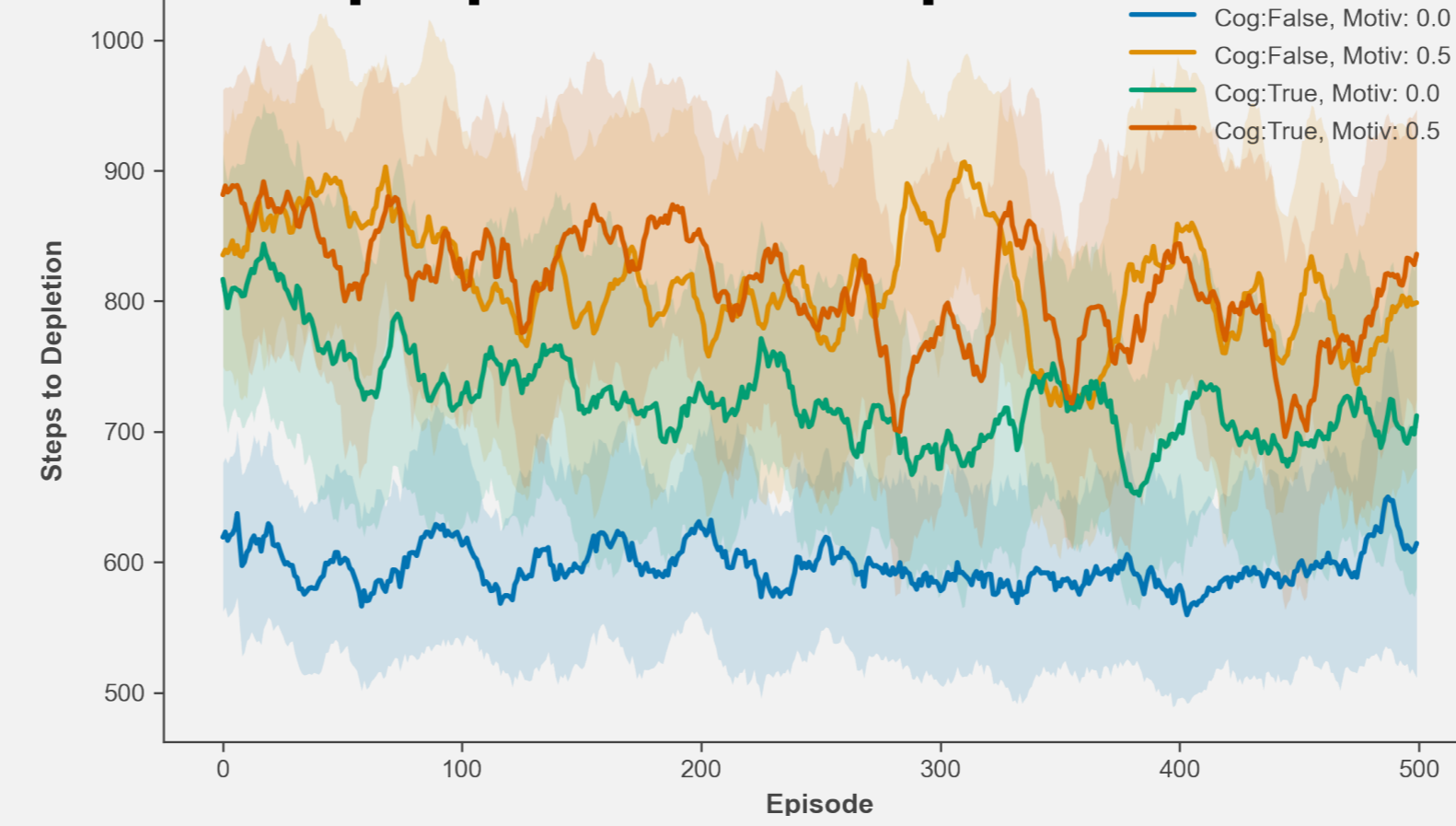
- 7 levels of empathy: $\alpha = \{0; 0.15; 0.25; 0.5; 0.75; 0.85; 0.99\}$
- All with the ability to observe the emotion of the other agents
- 24 simulations of 500 episodes each, 6 per condition

DEPENDENT VARIABLES

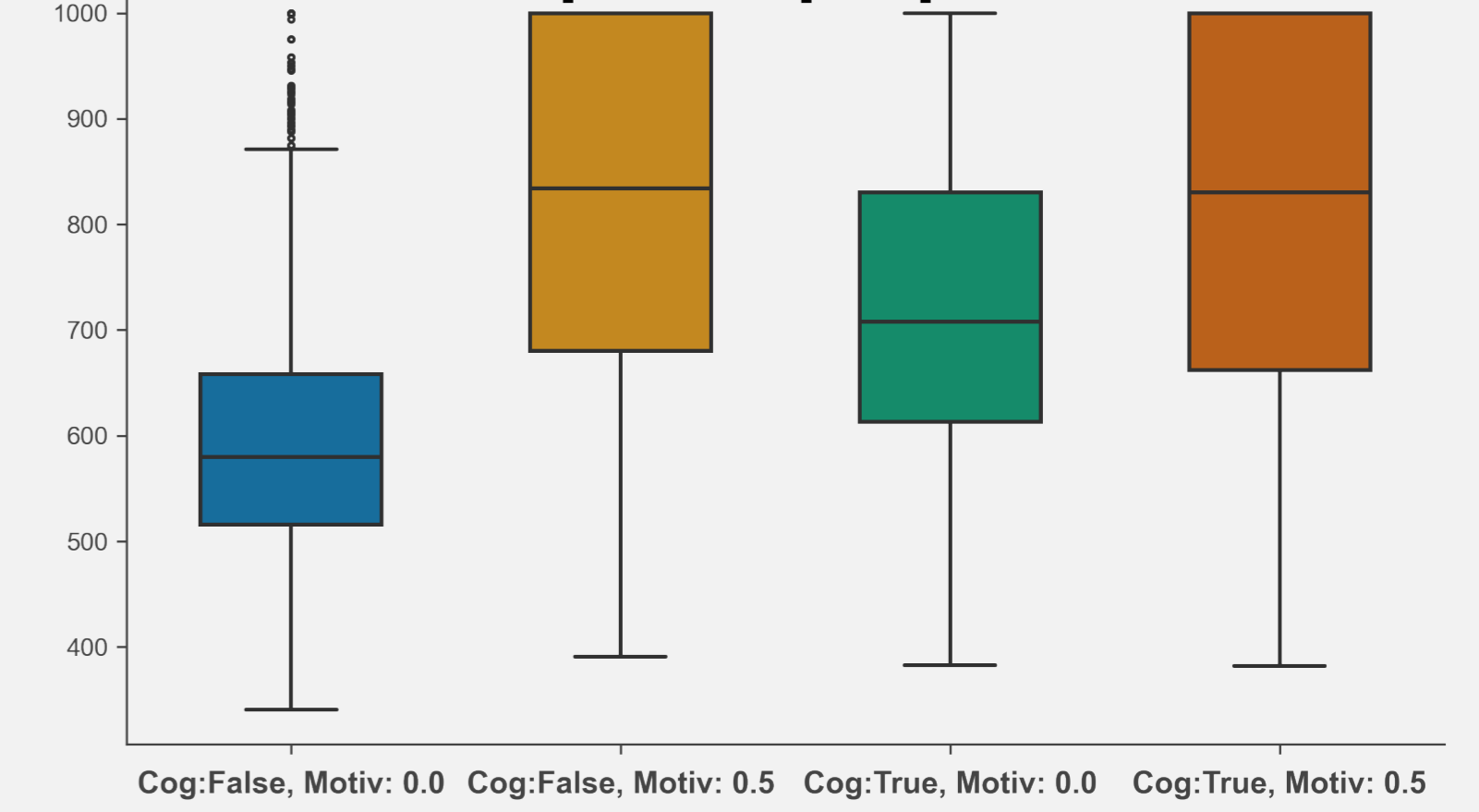
- Length of the episode**
Lower values indicate that the resources were depleted faster
- Gini coefficient**
Lower values indicate a more equal distribution of access to the resource

Study 1: empathy reduces overconsumption

Episode length as a function of the empathic properties and episodes

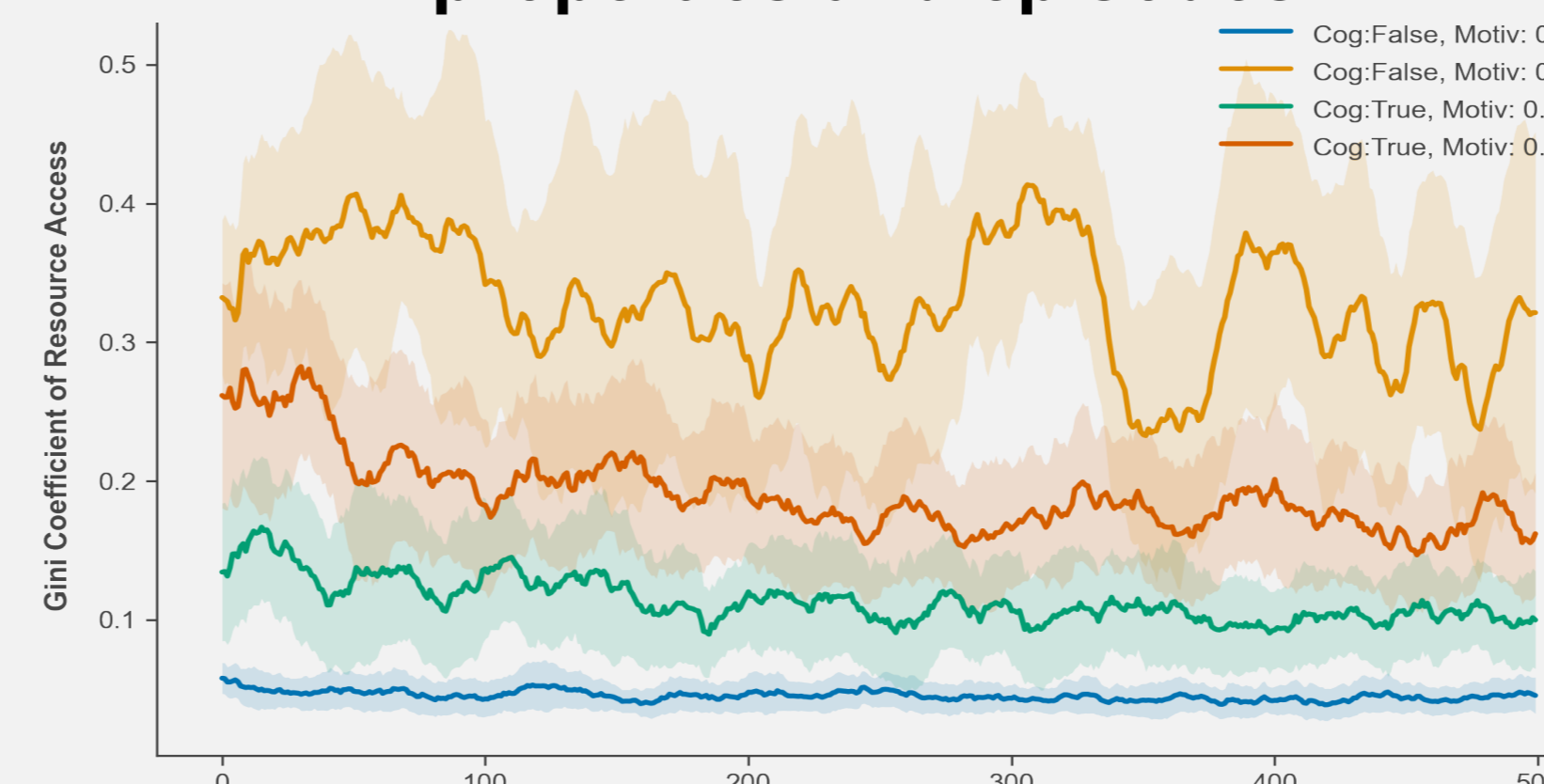


Episode length as a function of the empathic properties

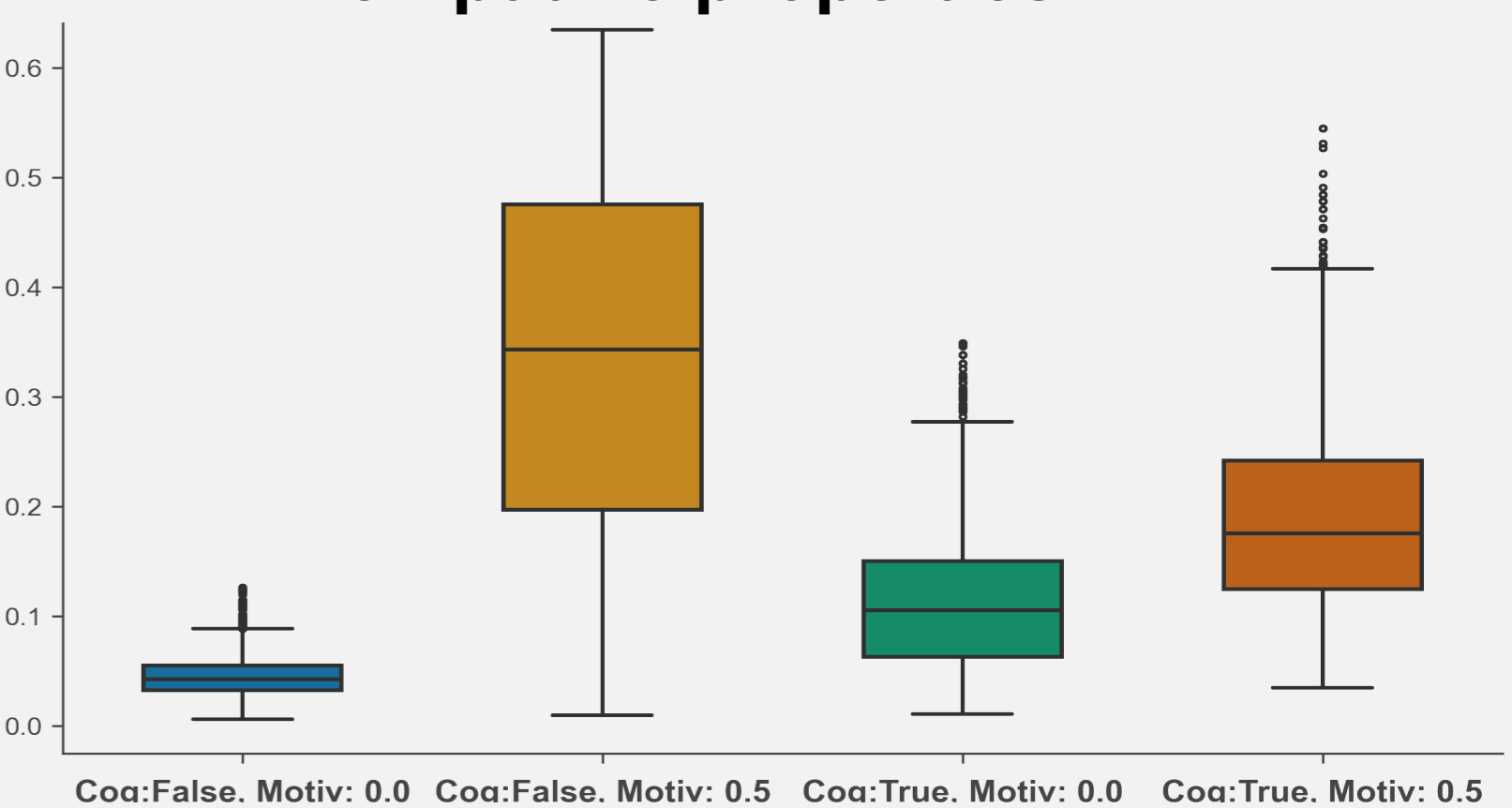


A two-way ANOVA revealed significant main effects of the ability to see emotions, $F(1, 8) = 68.29$, $p < .001$, $\eta^2 = 0.12$, and empathic motivation, $F(1, 8) = 396.53$, $p < .001$, $\eta^2 = 0.72$. There was also a significant interaction, $F(1, 8) = 75.53$, $p < .001$, $\eta^2 = 0.14$.

Fairness as a function of empathic properties and episodes



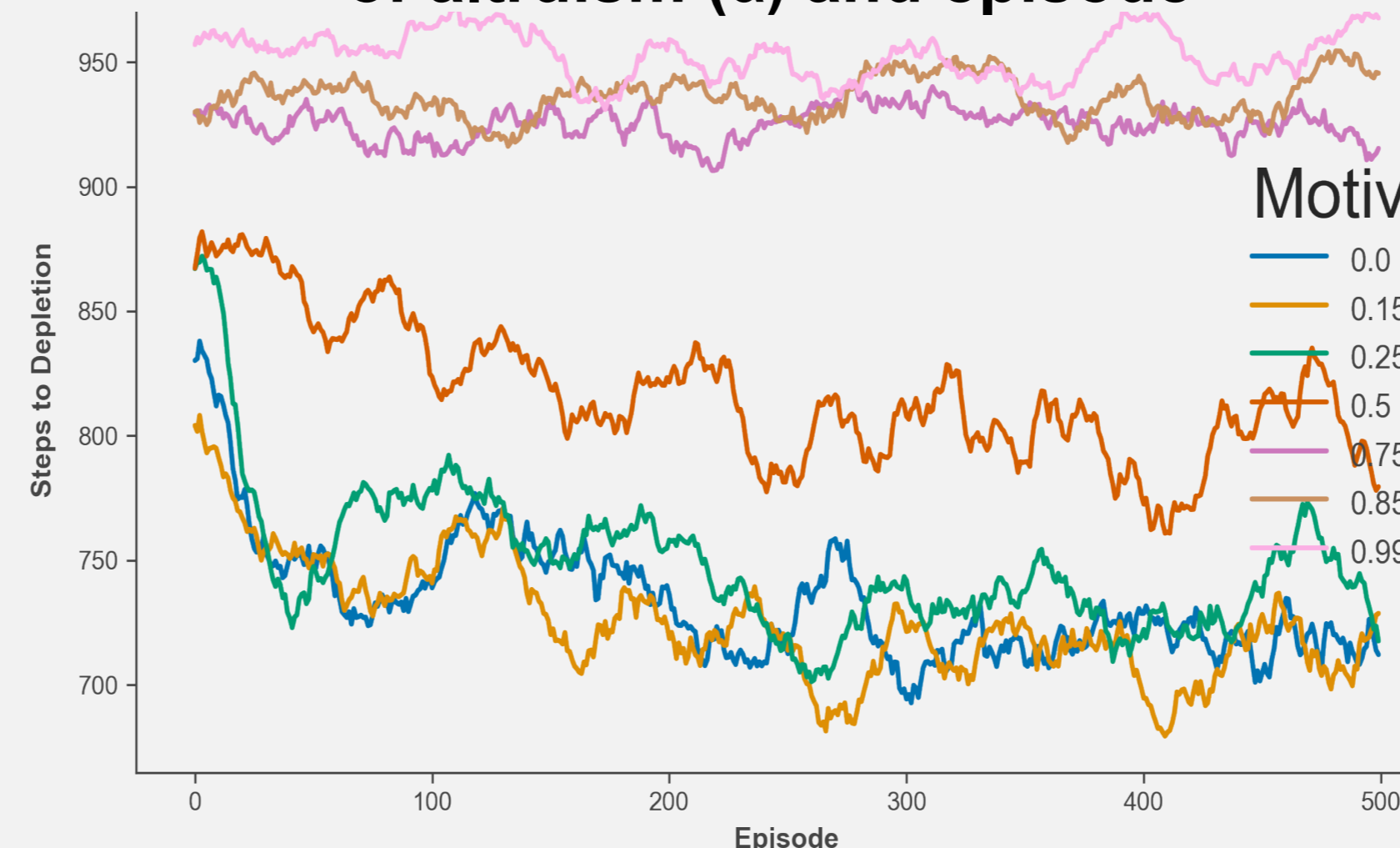
Fairness as a function of empathic properties.



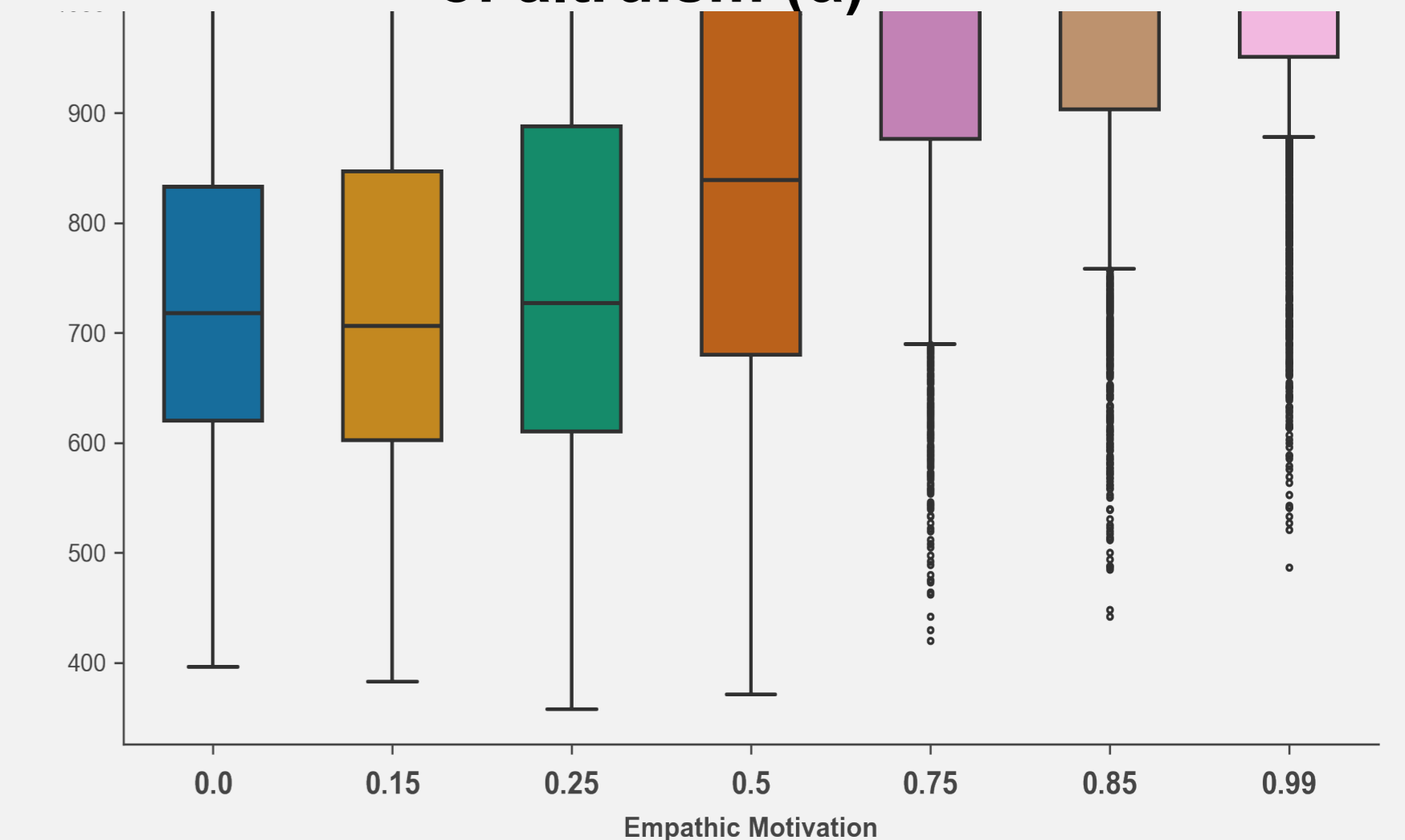
A two-way ANOVA revealed significant main effects of the ability to see emotions, $F(1, 8) = 68.29$, $p < .001$, $\eta^2 = 0.12$, and the empathic motivation, $F(1, 8) = 396.53$, $p < .001$, $\eta^2 = 0.72$. There was also a significant interaction $F(1, 8) = 75.53$, $p < .001$, $\eta^2 = 0.14$.

Study 2: more altruistic motivation, less depletion

Episode length as a function of the degree of altruism (α) and episode

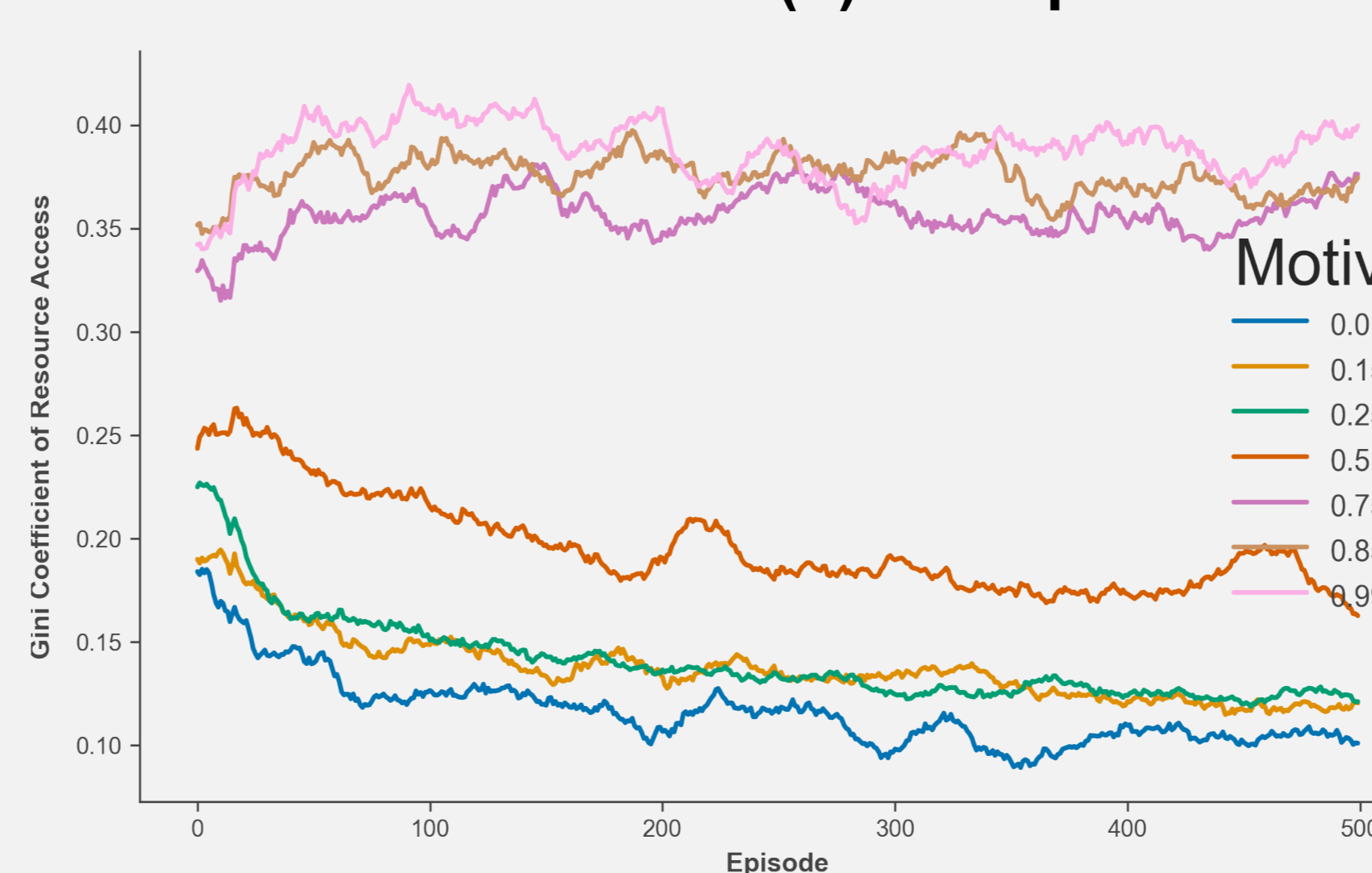


Episode length as a function of the degree of altruism (α)

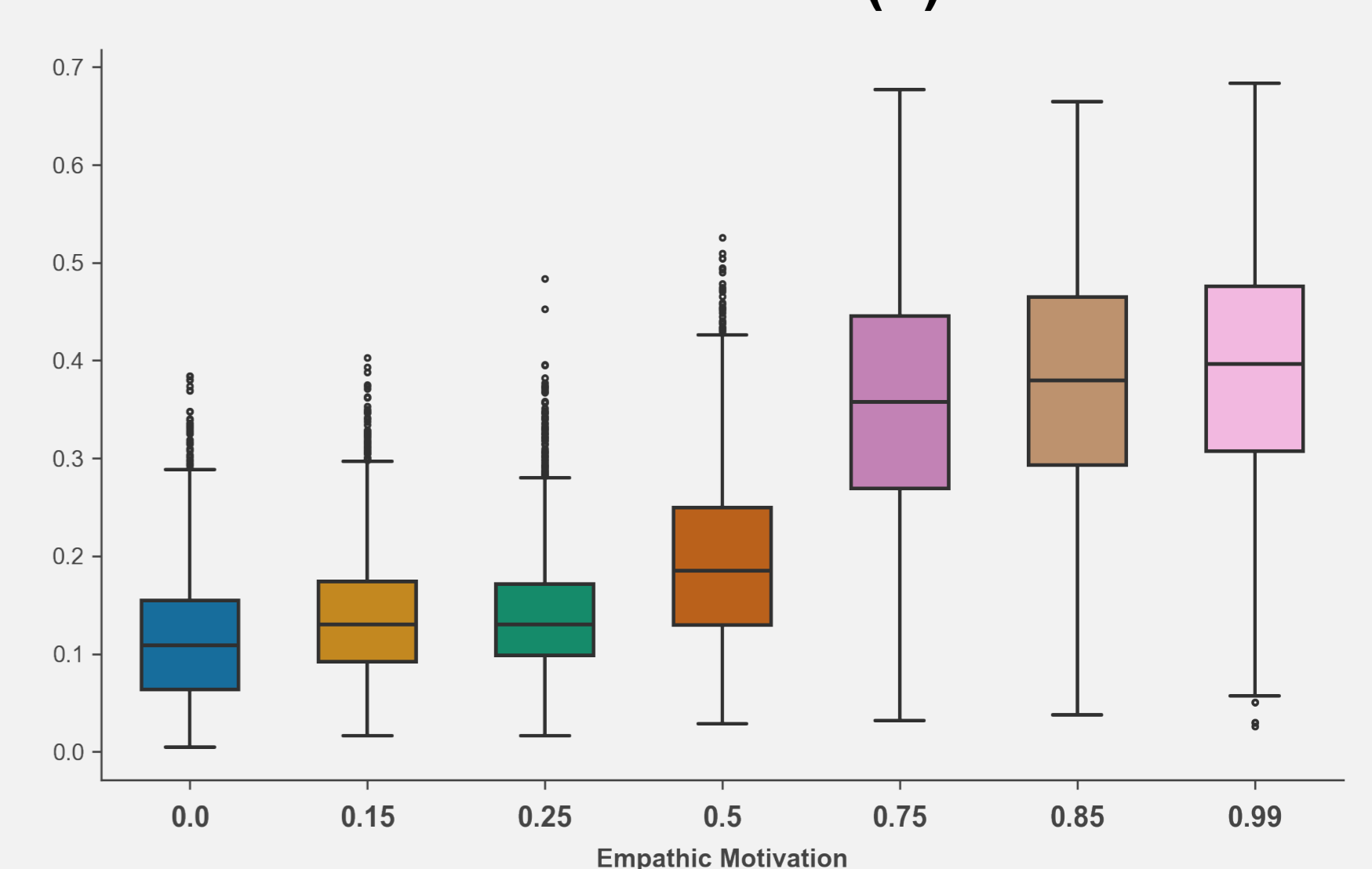


There is a positive Spearman correlation between the number of steps before depletion and the empathic motivation parameter ($r_s(22) = 0.92$, $p < 0.001$).

Fairness as a function of the degree of altruism (α) and episode



Fairness as a function of the degree of altruism (α)



There is a positive Spearman correlation between fairness and the empathic motivation parameter ($r_s(22) = 0.96$, $p < 0.001$).

Conclusion

Study 1

- Cognitive and motivational components influence behavior,
- empathic motivation is the main driver of sustainability.
- perceiving others' emotions increases fairness in resource distribution.

Study 2

Increasing the weight of others' rewards enhances sustainability but reduces fairness among agents.



Multi-agent reinforcement learning investigation of the mechanistic role of empathy on pro-environmental behavior

Iste Jérôme^{1,2,3,*}, Rabeau Evann^{1,2,*}, Ismael T. Freire³, Mehdi Khammassi³

¹Université Paris Cité, France; ²Sorbonne Université, France; ³ISIR – Institut des Systèmes Intelligents et de Robotiques

*Authors with equal participation



OPERATIONALIZATION OF EMPATHY
 * Cognitive component → agents observe emotional signals derived from others' recent consumption rates
 * Motivational Component → agents' reward is modulated by the average of others' emotional signals
 * R_i^{self} = agent's own consumption outcome
 * e_j = emotional signal of agent j , derived from their recent consumption rate (via a linear mapping)
 * N = number of agents
 * α = empathic weight parameter

Introduction

- The ecological crisis is commonly conceptualized as the Tragedy of the Commons (Hardin, 1968) where common pool resources risk being depleted because of individual's consumption.
- Empathy has been identified as a crucial factor, but the causal role of its mechanisms are difficult to test empirically at scale (Lynne et al. 2016). What is the role of the cognitive, affective and motivational component ?
- Prior work showed that modifications of agents' reward structures or observations can induce cooperation, conventions, and sustainability (Perolat et al., 2017; Wang et al., 2019; Zhu & Kirley, 2019). Using multi-agent reinforcement learning, we study the influence of the components of empathy (cognitive, affective, motivational) on common-pool resource dilemmas. We expect the conjunction of the ability to observe and be motivated by others' emotions to reduce overconsumption.

Methods

ENVIRONNEMENT

- Single renewable common-pool resource shared among 6 agents
- Resources regenerate dynamically based on the amount of resources left in the pool
- Episodes end at full depletion of the pool of resource or after a fixed number of steps (1000)
- Resource stock resets each episode, but agents retain learned policies

AGENTS

- All agents use Deep Q-Networks (DQN) to learn action-selection policies
- ϵ -greedy strategy: starts exploratory, then shifts toward exploitation
- Agents share identical network architectures and hyperparameters
- Manipulation of empathy level : α
- Manipulation of the ability to observe others emotions : yes (true) or no (false)

OPERATIONALIZATION OF EMPATHY

- Cognitive component → agents observe emotional signals derived from others' recent consumption rates
- Motivational Component → agents' reward is modulated by the average of others' emotional signals

$$R_i^{empathic} = (1 - \alpha) \cdot R_i^{self} + \alpha \cdot \frac{1}{N-1} \sum_{j \neq i} e_j$$

- R_i^{self} = agent's own consumption outcome
- e_j = emotional signal of agent j , derived from their recent consumption rate (via a linear mapping)
- N = number of agents
- α = empathic weight parameter

STUDY 1 : MULTIPLE ALPHA

- 7 levels of empathy
- All with the ability to observe the emotion of the other agents

STUDY 2 : EMPATHY COMPONENT MATRIX

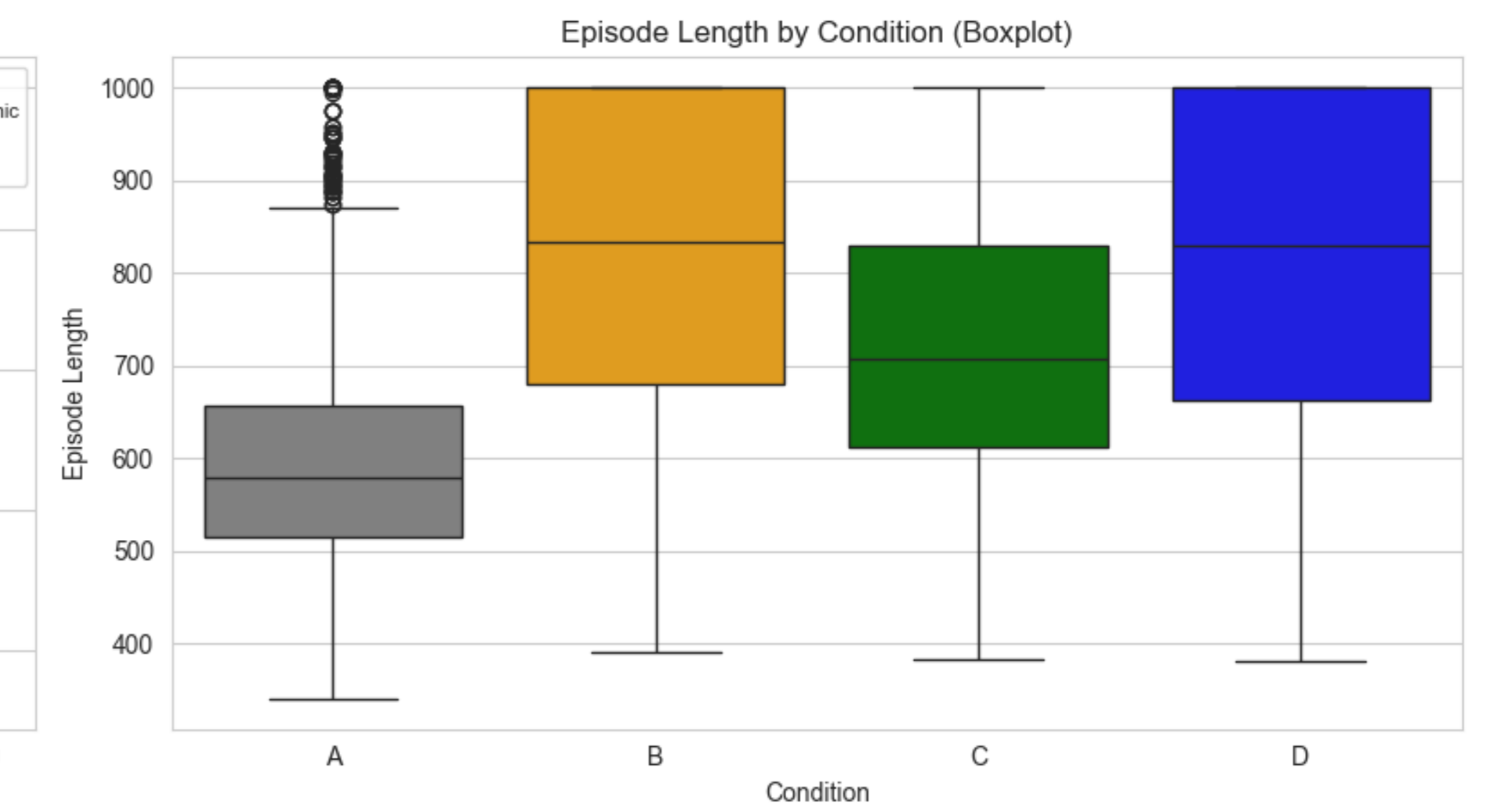
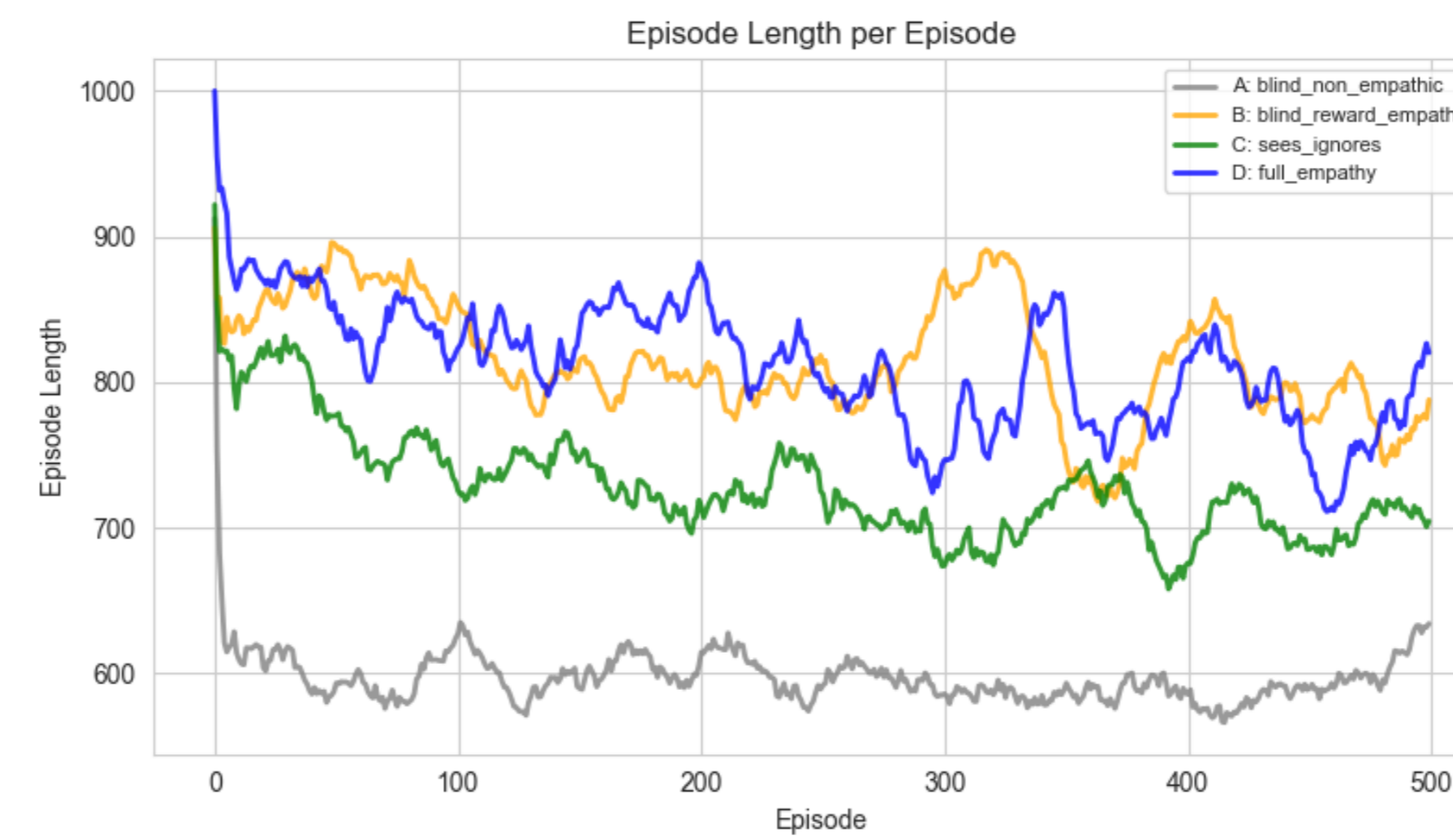
BLIND, NON EMPATHIC	BLIND, REWARD EMPATHIC
SEES emotion, ignores them	SEES emotion, Reward Empathic

DEPENDANT VARIABLES

- Length of the episode (maximum of a 1000 before the episode ends)
- Gini coefficient

Results study 1

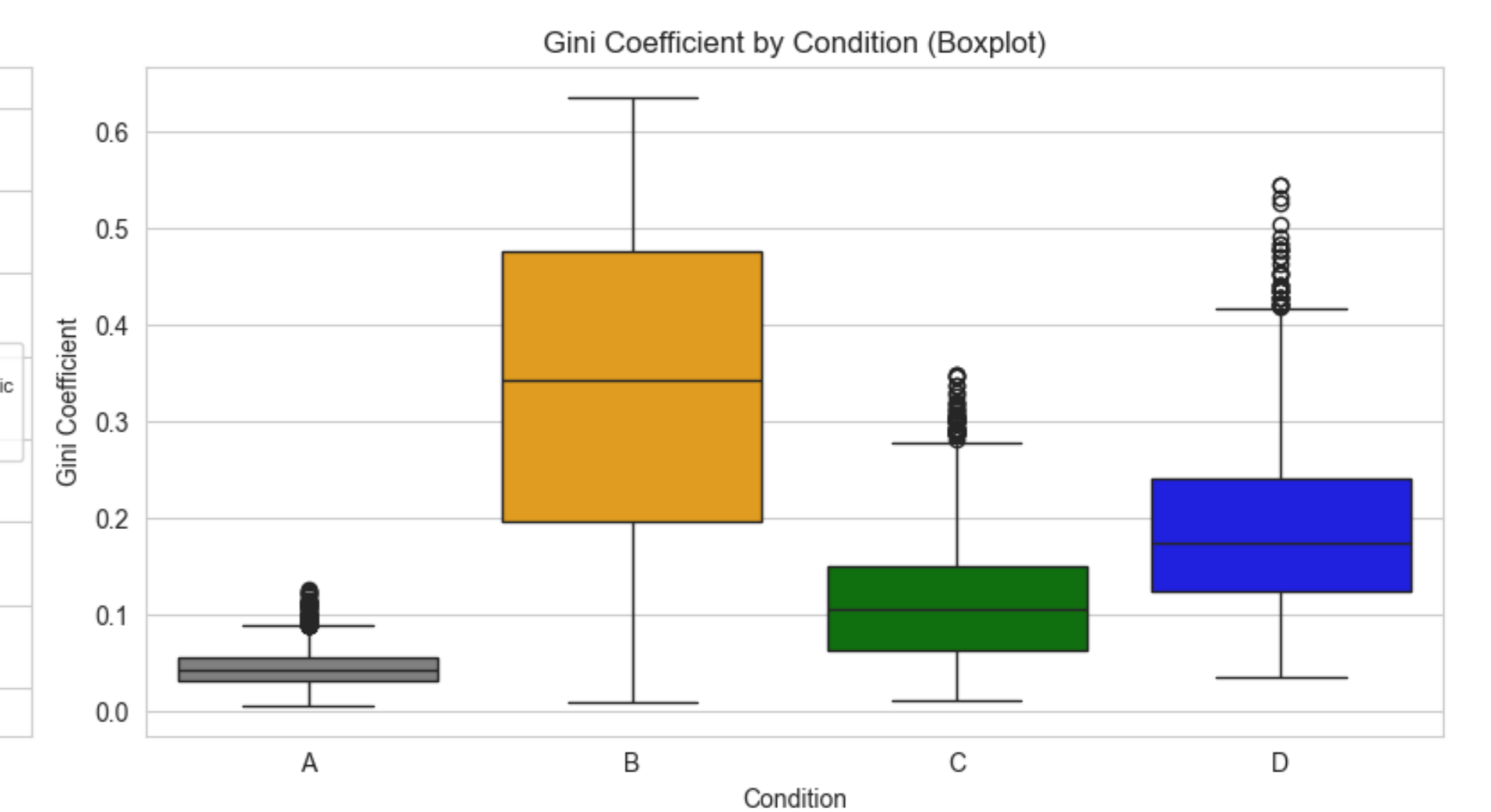
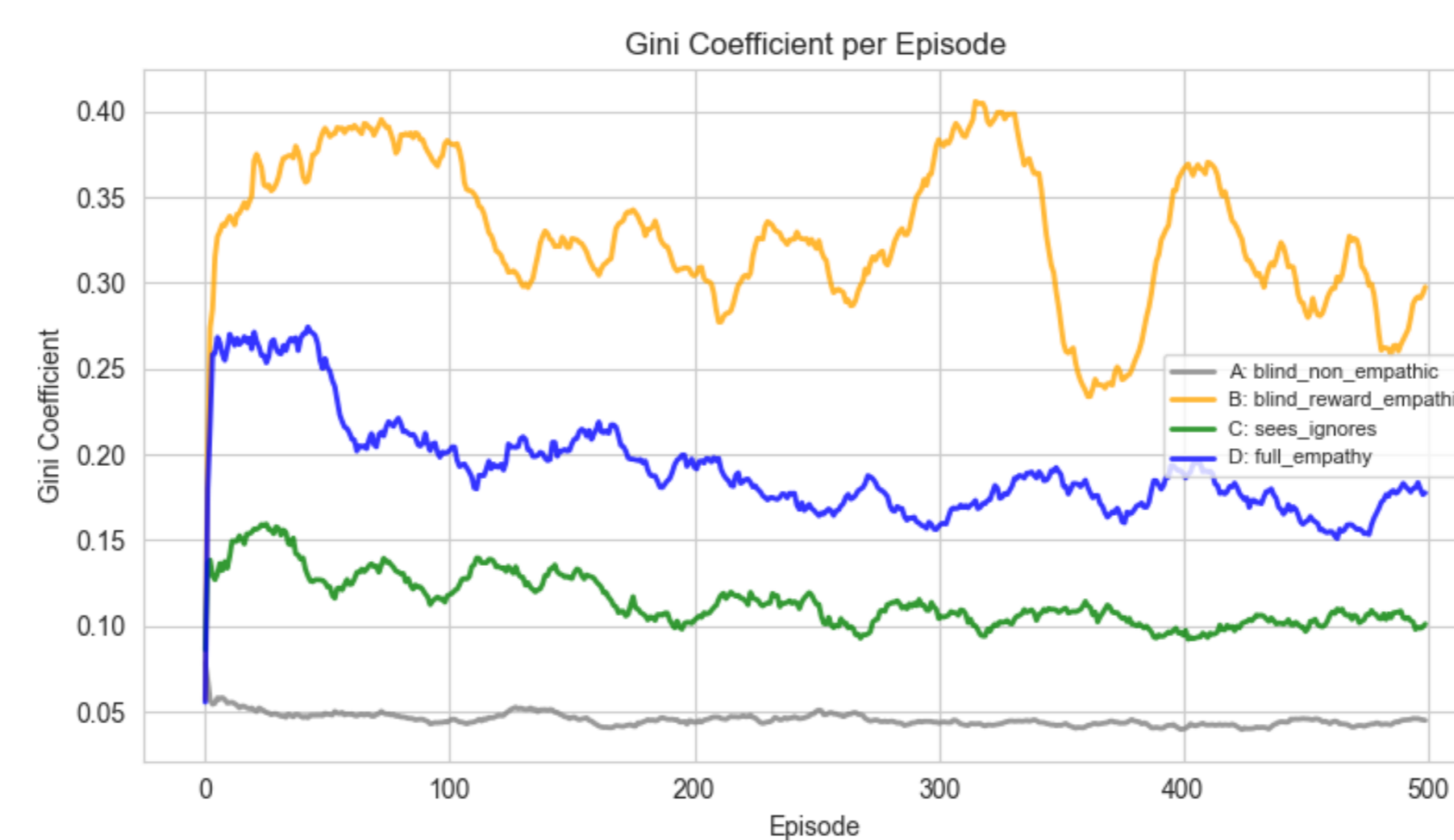
EPISODE LENGTH : TRAJECTORY AND DISTRIBUTION



2(see_emotion: True vs. False) x 2(alpha_level: 0 vs. 0.5)

Intracation : rapport
 Effet see_emotion
 Effet alpha_level

GINI COEFFICIENT : TRAJECTORY AND DISTRIBUTION

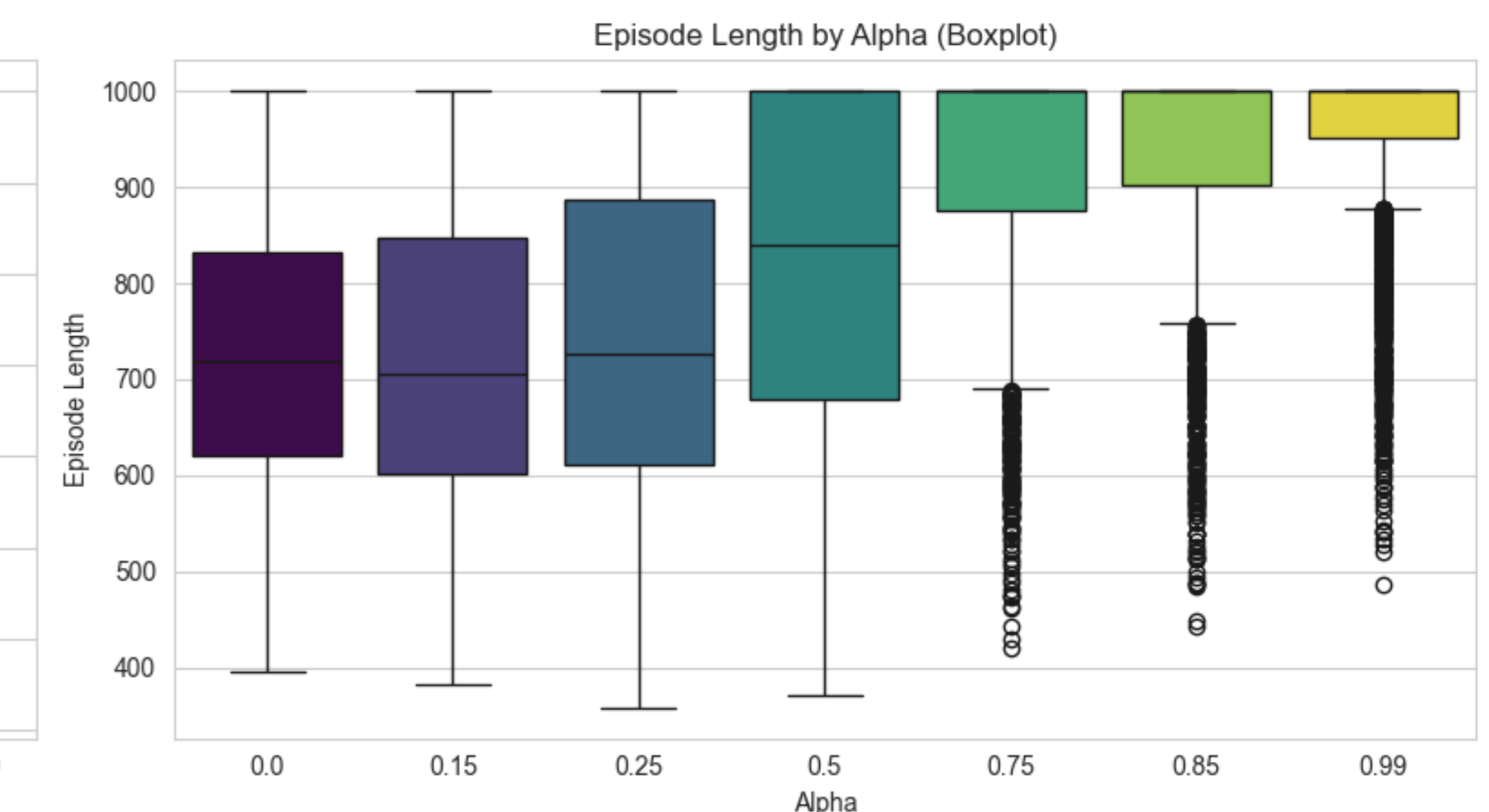
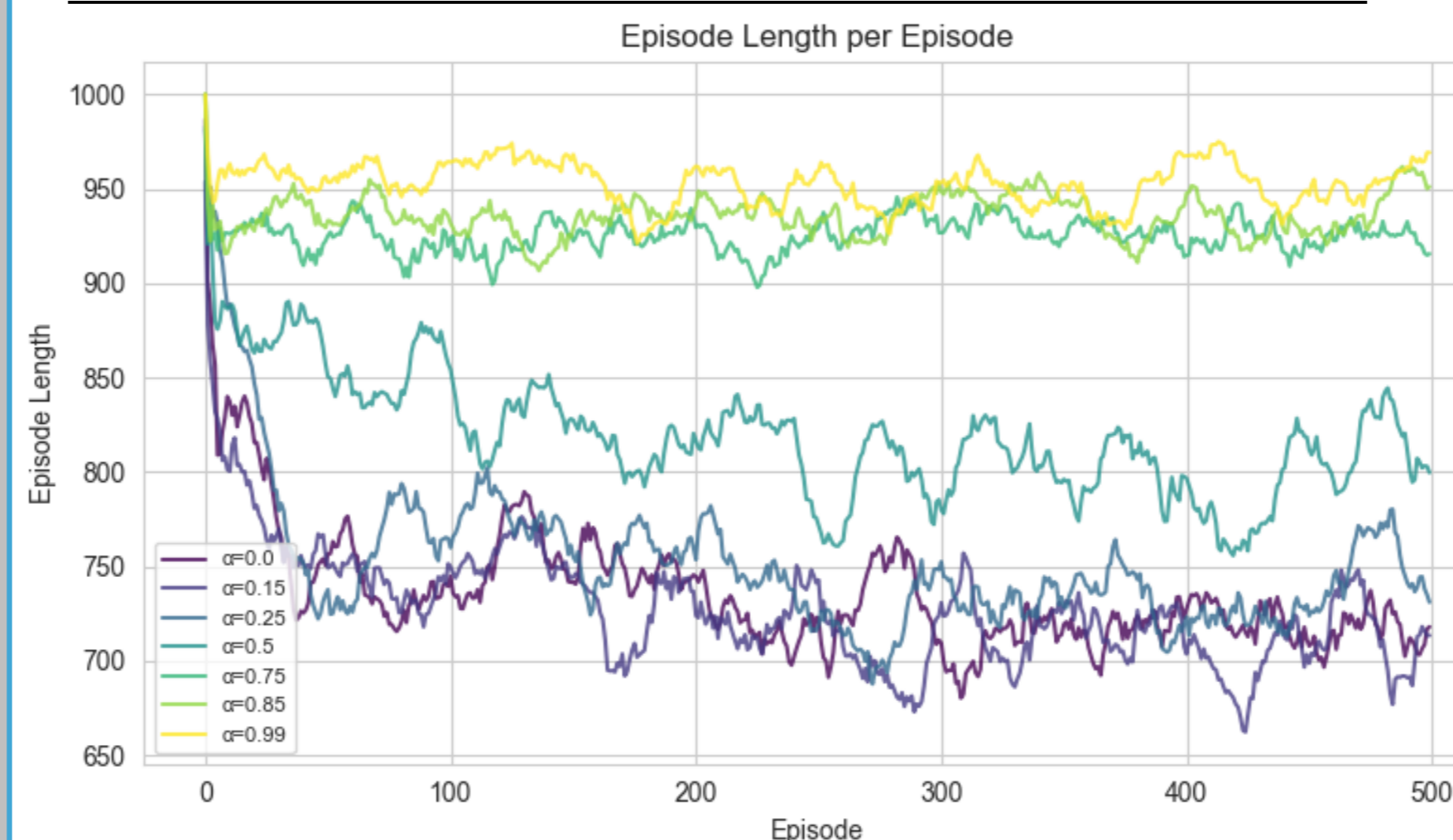


2(see_emotion: True vs. False) x 2(alpha_level: 0 vs. 0.5)

Intracation : rapport
 Effet see_emotion
 Effet alpha_level

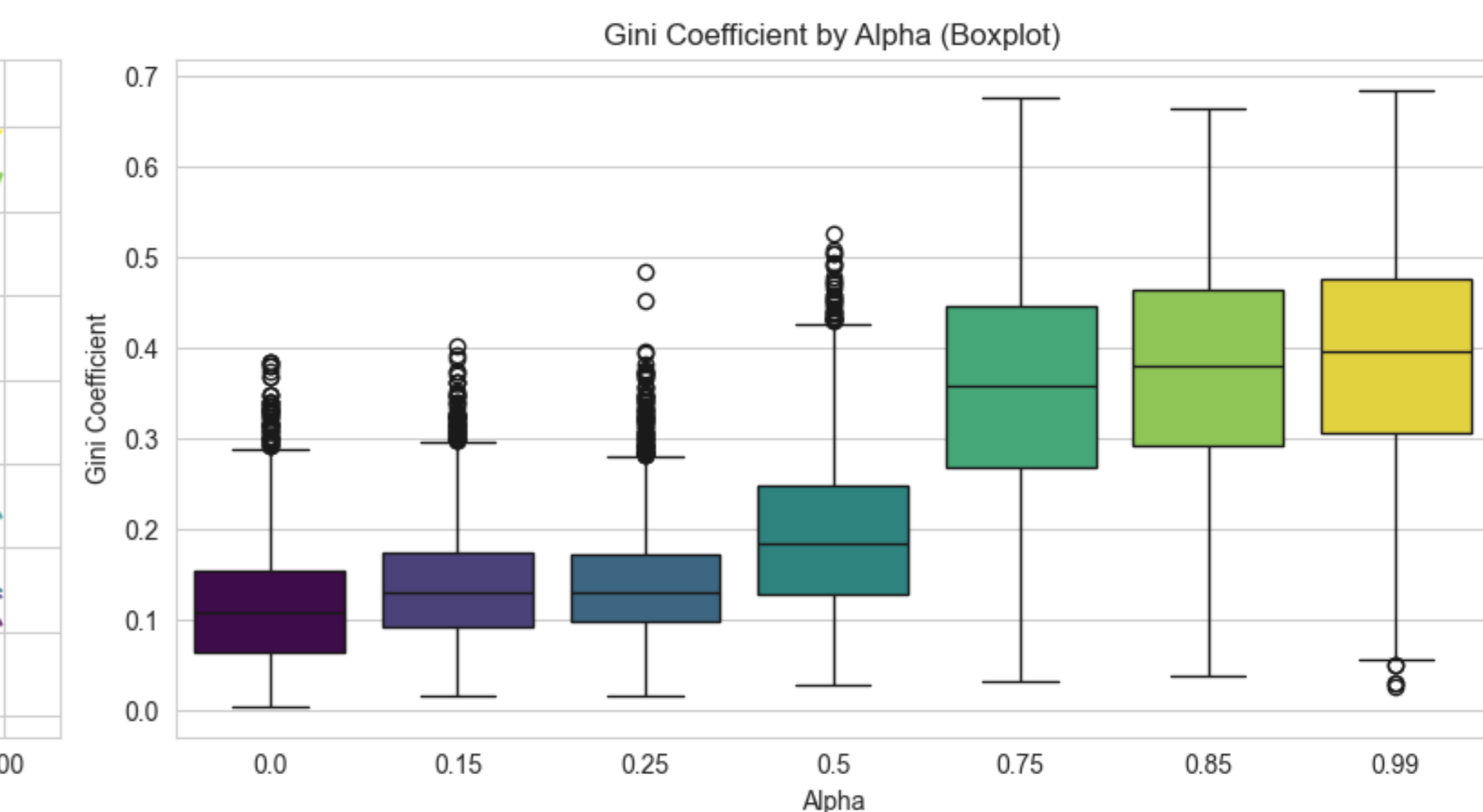
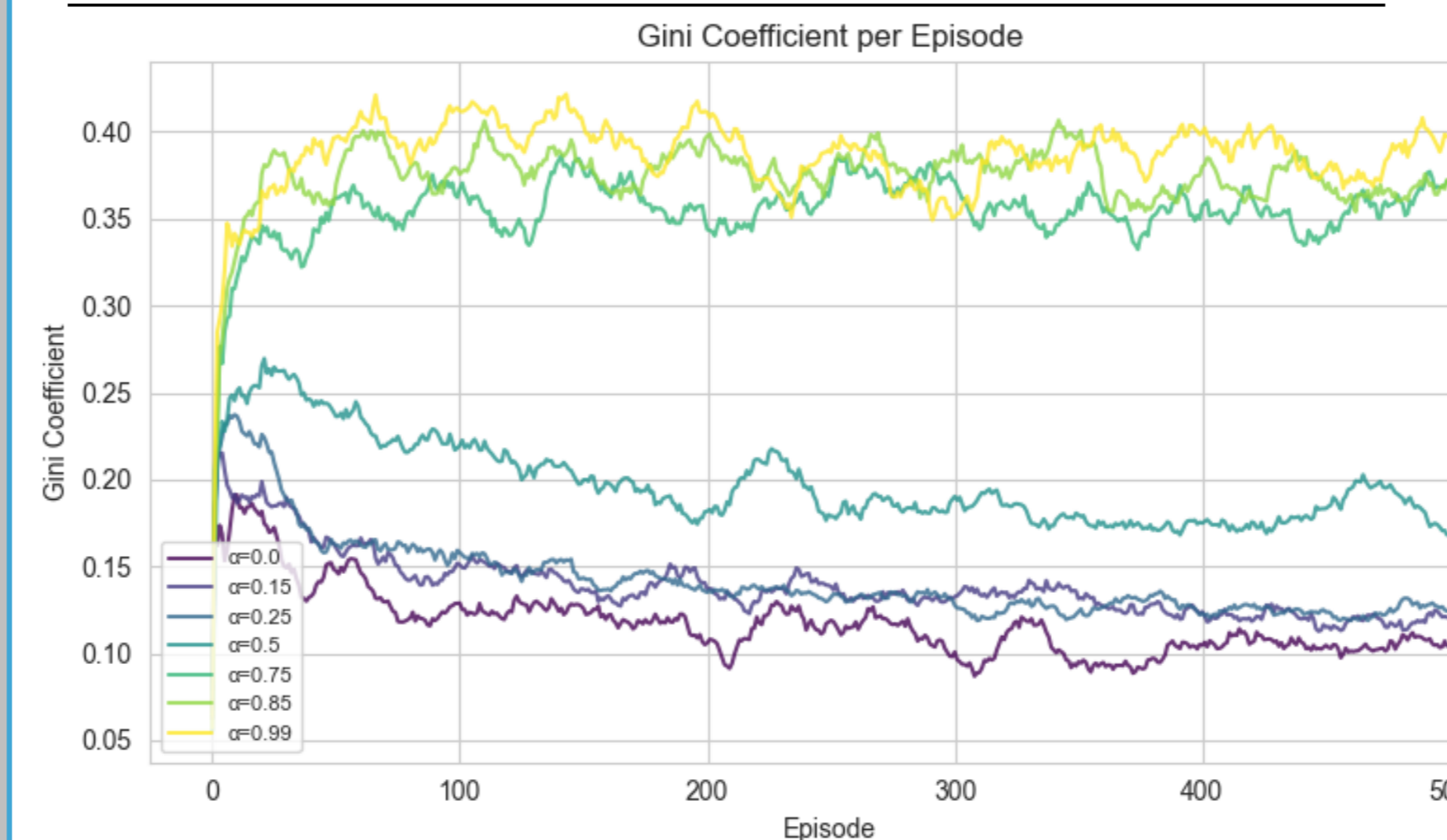
Results Study 2

EPISODE LENGTH : TRAJECTORY AND DISTRIBUTION



Linear regression : slope = 253, $R^2 = 0.93$, $p < 0.001^{***} < 0.0167$

GINI COEFFICIENT : TRAJECTORY AND DISTRIBUTION



Linear regression : slope = 0.31., $R^2 = 0.93$, $p < 0.001^{***} < 0.0167$

Conclusion

- Study 1 : Empathy has an effect on resource consumption and fairness. Both cognitive and motivational properties have an effect but the motivational factor appears to be a stronger driver of sustainability.
- Study 2 : The effect of the motivational factor is such that increasing others' utility in the learning process increases sustainability but also decreases the fairness of the repartition of resources among agents.
- Main take away : empathic concern might be the main driver of empathetic effects on pro-environmental behaviors. However, the ability to see emotions seems to increase fairness when agents are altruistically motivated.