

Science and Technology Group  
Annual Report FY 2023  
Xinying Jia  
Science and Technology Associate

## 1 Introduction

In my research, I intend to elucidate the dynamics of cooperation involving the social sentiments of fear, shame, and guilt. These three types of incentive cultures are well documented by cultural anthropology. To date, game-theoretical models addressing the social dynamics and evolution of incentive systems have largely restricted attention to the mechanisms associated with fear-based cooperation, ignoring the crucial roles played by other social sentiments. I investigate under which conditions fear-, shame-, and guilt-based cooperation are expected to emerge through cultural evolution and how transitions between and mixtures of these different incentive cultures are likely to occur. My study will contribute to the understanding of the complex social dynamics underlying fear-, shame-, and guilt-based cooperation.

The project is planned to proceed in three steps. First, reviewing the literature to identify empirical evidence and sociological assessments of the mechanisms enabling fear-, shame-, and guilt-based cooperation. This includes analyzing the relative strengths and weaknesses of these mechanisms, as well as the differential conditions under which they are likely to function, falter, or fail. Second, developing a series of innovative game-theoretical models that capture the key insights established by this review. These models will enable studying the complex social dynamics underlying fear-, shame-, and guilt-based cooperation. Third, considering how to integrate these models into an overarching framework that will allow me to investigate under which conditions fear-, shame-, and guilt-based cooperation are expected to emerge through cultural evolution and how transitions between and mixtures of these different incentive cultures are likely to occur. Particular attention will be given to the cultural evolution leading to the gradual internalization of incentive processes, as a society's dominant social sentiments shift from fear to shame and from shame to guilt.

## 2 Activities and Findings

I have worked intensively with an intern student Isabelle Oberlin on the computational experiments of the game theoretical model.

The method we use is to analyze social learning in a group of agents. In each round, all agents receive the same benefit and pay the same relative maintenance costs on their capitals. Some agents commit to a risky behavior providing them with an additional fixed benefit at the cost of exposing them to potential punishment occurring at random moments with given frequency and severity. Each agent can imitate the behavior of a randomly selected other agent with a probability determined by a sigmoid function of their current capital difference.

Our main findings are:

- Social learning through imitation allows inferior strategies to spread despite causing long-term losses.
- Inferior strategies spread because the dependence of imitation probabilities on gains and losses inevitably is nonlinear.
- Inferior strategies spread most easily when imitation probabilities sensitively depend on capital differences and the latter's distribution is wide and left-skewed.
- Effective deterrence against socially deviant strategies requires higher punishment frequencies and severities than suggested by assessing long-term losses.

In summary, when punishment is severe but rare, the committing agents are more likely to be imitated than the non-committing agents, even though the average capital of the committing agents is much lower than that of the non-committing agents.

The first draft of the paper has been accomplished.

### 3 Collaborations

Collaborator: Ulf Dieckmann, OIST; Isabelle Oberlin, Intern student of OIST

### 4 Output

Poster presentation of the project at the STG forum in March 2024.