

Observation Enhances Third-Party Punishment Only among People Who Were not Hot-Tempered

Yutaka Horita^{1,2}, Masanori Takezawa^{3,*}

1. National Institution of Informatics, 101-8430, Japan

2. JST, ERATO, Kawarabayashi Large Graph Project, 101-8430, Japan

3. Hokkaido University, 060-0810, Japan

* Author for correspondence (m.takezawa@let.hokudai.ac.jp)

The evolution of punishment toward norm-violators has been discussed for understanding a large-scale human cooperation. Recent studies showed that the presence of cues of surveillance makes people concern about their reputation and increase altruistic behavior. Recent study also suggests that explicit cues of observation affect punitive behavior. We examined whether both explicit (being observed by an experimenter) and implicit cues (drawing of stylized eyes) of observation enhance third-party punishment. The results of the experiments with Japanese participants showed that both type of cues of observation increased third-party punishment only among those who did not feel anger toward an unfair allocator. In contrast, the cues suppressed the punishment when participant felt stronger anger toward an allocator. Our study suggests that we humans are endowed with the psychological system inducing third-party punishment in response to cues of being observed while its function may be moderated by cultural factors.

Keywords

punishment, reputation, third-party punishment, emotion

Introduction

Punishment toward norm-violators has been considered to be a key for sustaining a large-scale cooperation. It has been well known that people incur costs of punishing norm-violators even though they themselves did not suffer from norm violation. Such punitive behavior is called third-party punishment (Fehr & Fischbacher, 2004).

However, how did costly punishment itself evolve? If punishers receive more personal benefits for their punitive behavior from other group members, punishers will evolve through a process of natural selection. Indirect reciprocity model suggested that costly altruistic behavior becomes beneficial, if people selectively give resources to those who have a good reputation (Nowak & Sigmund, 1998).

Many empirical studies demonstrated that people had a psychological mechanism to be sensitive to the cues of reputation. In a series of these studies, Haley and Fessler (2005) first showed that even subtle cues of surveillance, such as drawing of stylized eyes enhanced altruistic behavior (see also, Nettle, et al., 2013). Previous studies also found that people gave punishers a good reputation as a trustworthy person more than non-punishers (Barclay, 2006; Horita, 2010). If punishment toward a norm-violator results in acquiring a good reputation, people will be more likely to engage in punishment when cues of observation are presented. Kurzban, DeScioli, and O'Brien (2007) found that participants were willing to engage more greatly in third-party punishment when their behavior would be explicitly known to other participants or an experimenter. This suggested that costly punishment would also be promoted by concern for acquiring a good reputation.

The purpose of this study is to investigate the effects of observation on promoting costly punishment. Studies that have investigated whether the cues of observation affect third-party punishment are still not enough and only the role of explicit cues of observation was investigated so far. In this study, we examined effects of both explicit and implicit cues on costly punishment in the third-party punishment game.

Methods

Participants

Sixty-six undergraduate students (22 males and 44 females; mean age = 19.5 ± 1.2) participated at Sophia University, Japan. Monetary rewards were emphasized for their participation.

Third-party punishment game

Participants played the one-shot third-party punishment game. There were 3 roles in the game; an allocator, a recipient, and a third-party (In the instruction, these roles were called as X, Y, and Z, respectively). First, an allocator decides how to divide 1,000 yen (about US \$10) between him/herself and a recipient by choosing one of two options: a fair (500 yen to the allocator, 500 yen to the recipient) or an unfair option (900 yen to the allocator, 100 yen to the recipient). The recipient receives the amount of money allocated by the allocator. After the allocation, a third-party is informed of the option the allocator has chosen. The third-party receives 500 yen from the experimenter and is given an opportunity to reduce the allocator's payoff by paying a cost from his/her amount of money. The third-party can spend a maximum of 300 yen in a unit of one yen. Tripled amount of money that the third-party has paid is reduced from the allocator's payoff.

All participants were assigned to the role of a third-party, and were informed that an allocator chose the unfair option. Participants then decided how much money they spent to reduce the allocator's payoff.

Conditions

The type of cues of observation was manipulated as a between-participant factor: the Control ($n = 22$), the Eye ($n = 23$), and the Experimenter condition ($n = 21$). Participants were assigned to one of the three conditions randomly.

Procedures

When participants arrived in the laboratory, they received an ID number from a receptionist. Participants were informed that each participant would be placed in a separate room. Upon entering the room, participants received 200 yen as a show-up fee. Participants received printed instructions of the third-party punishment game. It was emphasized that all the participants were never faced other participants during and after the experiment. After reading the instructions, participants were asked to fill out a few questions to confirm whether they understood the instructions. After all participants completed the questions correctly, they were told that the rest of all the instructions would be displayed on a computer screen.

In the room, there was a laptop computer on the desk. During the experiment, abstract drawing was displayed in the Control and Experimenter conditions and a drawing of stylized eyes was displayed in the Eyes condition. These drawings were the background of experimental program (see, Figure 1).

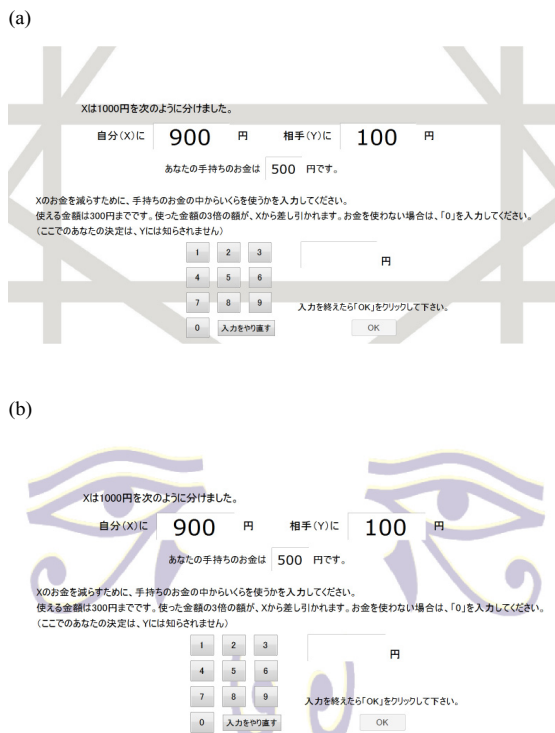


Figure 1. Decision screen displayed on the computer screen in (a) the Control and the Experimenter conditions and (b) the Eye condition.

In the Control and Eyes conditions, participants were left alone in the room. Participants in these conditions were explained that their decisions were delivered to an assistant, who never saw them, via computer for calculating an amount of rewards. Thus, in both conditions, anonymity of participants was assured. In the Experimenter condition, an experimenter was staying in

the room during the experiment. The experimenter sat diagonally to the front of participants with another laptop computer connected to the participant's computer with a cable. Participants were explained that their decision whether to punish an allocator would be displayed on the experimenter's computer screen for calculating an amount of rewards.

After the participants read printed instructions, the computer program instructed that participants were assigned to the role of the third-party. Subsequent instructions followed that the other participants who were assigned to the role of the allocator would decide how to split 1,000 yen at first, and then the allocator's decision would be displayed on the third-parties' computer screen. After a few minutes, they were informed that the allocator chose an unfair option. Participants were then asked to input the amount of money they wanted to pay to reduce the allocator's reward.

After participants completed their decision, the drawing on the computer screen disappeared. Then, they were asked to fill out a post-experimental questionnaire. After participants finished filling out it, they were paid the amount according to their decision and additional 200 yen as a show-up fee. The experiment took about an hour. Participants were debriefed after all experimental sessions finished.

Results

We excluded one participant whose nationality is China from the following analysis in order to control the effect of participants' nationality.

Punishment

The mean amount of money spent on punishment was 73.8 yen ($SD = 102.0$) in the Control condition, 73.9 yen ($SD = 104.3$) in the Eye condition, and 70.0 yen ($SD = 99.5$) in the Experimenter condition, respectively. There were no significant differences across three conditions (A one-way ANOVA; $F(2, 62) = 0.01, p = .99$). The proportion of participants who spent money to punish the unfair allocator also did not differ across three conditions (38.1 % in the Control condition, 43.5 % in the Eye condition, and 38.1 % in the Experimenter condition, respectively; $\chi^2(2) = 0.18, p = .91$).

The relationship between punishment and anger toward the allocator

Participants were asked how much anger they felt toward the allocator in the post-experimental questionnaire (1 = not at all, 7 = very strongly). Mean ratings of anger did not differ between the three conditions ($F(2, 62) = 0.97, p = .39$), that is, the unfair allocator induced the same degree of emotional responses in all the conditions. However, the feeling of anger toward the unfair allocator was significantly correlated with the level of punishment only in the Control condition ($\rho = .53, n = 21, p = .01$), but not in the Eye ($\rho = .00, n = 23, p = .99$) and the Experimenter conditions ($\rho = -.25, n = 21, p = .28$). This suggests that feeling of anger induced punishment only when participants were in anonymous settings and those who felt strong anger would not be always willing to punish the unfair allocator when implicit or explicit cues

of observation existed.

We divided participants into two groups by the mean level of anger ($M = 3.57$) and compared the amount of money spent for punishment. Figure 2 shows the means amount of punishment by conditions and the level of participants' felt anger, respectively. A two-way ANOVA revealed that the interaction effect of conditions and the level of anger was significant ($F(2, 59) = 4.74, p = .01$) and both the main effect of conditions ($F(2, 59) = 0.06, p = .94$) and the main effect of the level of anger ($F(1, 59) = 0.00, p = .96$) were not significant.

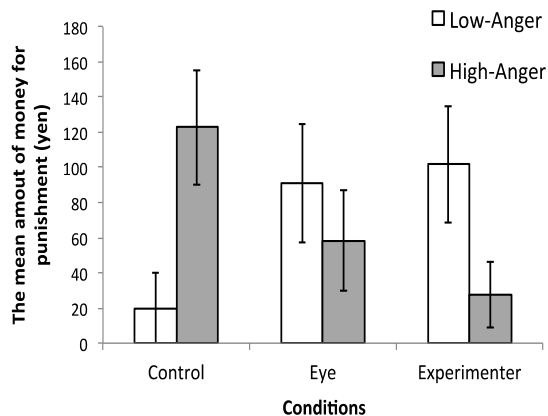


Figure 2. The mean amount of money paid for punishment by conditions and the level of participants' feeling of anger. Error bars indicate standard errors.

We conducted planned comparisons to test two types of observation effects; the Control vs. the Eye and the Experimenter (the effect of observation) and the Eye vs. the Experimenter (implicit vs. explicit). We revealed that participants who felt stronger anger spent more money for punishment in the Control condition than those in the Eye and the Experimenter conditions ($t(29) = 2.29, p = .03$). In contrast, among participants who felt lower anger, the amount of punishment in the Control condition were lower than those in the Eye and the Experimenter conditions ($t(28.1) = 2.47, p = .02$). These results suggest that high-anger participants punished less than when there was no observation, but low-anger participants punished more than when the observation was absent. Second, we tested the effect of types of cues on punishment. Among both types of participants, there was no significant difference between the Eye and the Experimenter conditions (high-anger participants, $t(29) = 0.74, p = .46$; low-anger participants, $t(20.9) = 0.23, p = .82$).

Discussion

We found that there was no difference in both the amount and frequency of punishment across the three conditions. These experimental results seem to contradict that of Kurzban et al. (2007) that found an explicit cue of being observed increased the amount of third-party punishment. However, we found that both explicit and implicit cues of being observed enhanced the amount of punishment among those who felt less anger toward an unfair allocator. In contrast, cues of being observed suppressed punishment

among those who felt strong anger.

One possible explanation for these findings may be attributed to a cultural difference in evaluation about punishment. Horita (2010) indicated that punishers acquired not only positive but also negative reputations. That is, punishers were not rewarded by others, while they were perceived as a trustworthy person. Punishment can also be perceived as aggressive behavior. Our participants who felt anger may have suppressed punishment when they faced the cues of observation in order not to spread their negative reputation as a hot-tempered person. On the other hand, those who did not feel anger toward an unfair allocator may have been influenced by the cues of observation to maintain their positive reputations because punishment was not a mean of negative emotion expression for them. As a consequence, the finding of Kurzban et al. (2007) could have been replicated only among our participants who felt less anger. Our findings suggest that there is the cultural difference in whether people evaluate punishment as being normative or aggressive and it may also play an important role.

Concerning the cooperation, evidence of the effects of cues of surveillance is rather mixed. Some studies reported that drawing of stylized eyes did not enhance cooperative behavior (Fehr & Schneider, 2009; Tane & Takezawa, 2011). Recent study also reported that longer exposure of cues did not have an effect to increase the level of altruistic behavior (Sparks & Barclay, 2013), so we should consider some effect of experimental settings. We showed that even explicit cues of observation did not have a direct effect on promoting punishment in contrast to previous study. Including the role of emotion or cultural backgrounds, exploring the conditions suppressing or enhancing the effects of cues of surveillance on cooperation and punishment is necessary to understand the evolution of cooperation.

Acknowledgement

This study was supported by Grant-in-Aid to Masanori Takezawa (23683011) and Yutaka Horita (11J02670) from the Japan Society for the Promotion of Science. We thank Genta Suzuki for his assistance in conducting the experiment and colleagues at Sophia University for helping recruitment of potential participants from their classes.

References

- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, 27, 325–344. (doi: 10.1016/j.evolhumbehav.2006.01.003)
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25, 63–87. (doi: 10.1016/S1090-5138(04)00005-4)
- Fehr, E., & Schneider, F. (2009). Eyes are on us, but nobody cares: Are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, 277, 1315–1323. (doi: 10.1098/rspb.2009.1900)
- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26, 245–256. (doi: 10.1016/j.evolhumbehav.2005.01.002)

- Horita, Y. (2010). Punishers may be chosen as providers but not as recipients. *Letters on Evolutionary Behavioral Science*, 1, 6–9. (doi: 10.5178/lebs.2010.2)
- Kurzban, R., DeScioli, P., & O'Brien, E. (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior*, 28, 75–84. (doi: 10.1016/j.evolhumbehav.2006.06.001)
- Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the dictator game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior*, 34, 35–40. (doi: 10.1016/j.evolhumbehav.2012.08.004)
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393, 573–577. (doi: 10.1038/31225)
- Sparks, A., & Barclay, P. (2013). Eye images increase generosity, but not for long: The limited effect of a false cue. *Evolution and Human Behavior*, 34, 317–322. (doi: 10.1016/j.evolhumbehav.2013.05.001)
- Tane, K., & Takezawa, M. (2011). Perception of human face does not induce cooperation in darkness. *Letters on Evolutionary Behavioral Science*, 2, 24–27. (doi: 10.5178/lebs.2011.15)