**Supplementary file 1**

**Methods**

**Design and Procedure**

A statistical priori-power analysis using G\*Power (Faul, Erdfelder, Lang, & Buchner, 2007) required 32 participants in the 2 (gender) × 2 (threat condition) × 5 (knowledge) experimental design (alpha prob = 0.05, 1-beta prob = 0.80, Number of groups = 4, Number of measurements = 5, Corr among rep measures = 0.5, Nonsphericity correction eta = 1.0). In line with previous studies using the same experimental paradigm (e.g., Yamagishi & Mifune, 2009), the effect size of interaction effects between gender and knowledge conditions was relatively weak. The 0.15 effect size revealed that the total sample size required was 160.

This study was carried out in accordance with the recommendations of the ethics committee of the Faculty of Humanities and Human Science in Hiroshima Shudo University, who approved the protocol for the study. All participants provided written informed consent in accordance with the Declaration of Helsinki.

Participants were 167 undergraduate students (90 women and 77 men) at Kobe University, Japan. They were randomly assigned to either the control or the outgroup threat condition. In each experimental session, four or five students were divided into two groups with a varied ratio of men and women within each group, but avoiding groups comprising only one gender. In eight of thirty-three sessions where only three participants were available, a confederate was added to the group and engaged into the same tasks as participants did

Participants were seated in a private booth and provided their informed consent. The experimental session comprised two ostensibly unrelated experiments in sequential order. First, the casually dressed experimenter in the first experiment asked participants to engage in a so-called “language task,” which was the prime for outgroup threat (Sugiura et al., 2017). The task required participants to search for and circle all nouns in three essays within five minutes. In the outgroup threat condition, the second essay was supposedly written by a person from another culture who was criticizing the attitude of the participants’ own country regarding a territorial issue (i.e., using the word “war”). In the control condition, the theme of the second essay was art. Participants then engaged in a simple calculation task for 30 seconds that served as a distraction.

After the first experiment, a new experimenter wearing a white lab coat entered the room and initiated the second experiment. Participants were asked to complete two decision-making tasks. In the first task, participants were asked about their preference between 28 pairs of abstract paintings. Participants were divided into either a “Klee group” or a “Kandinsky group,” based on the number of paintings by Klee or Kandinsky they chose. Based on total number Klee’s pictures chosen, half the participants in the session were assigned to the “Klee group” and others to the “Kandinsky group.”

In the second task, participants played the PD game. First, participants received and read instructions that included the following procedures. Participants were asked to engage in several rounds (the exact number was unspecified) of “transactions” with other participants. Their transaction partner would change each round. Participants were given 200 yen at the beginning of each transaction and decided how much of the 200 yen they would provide to their transaction partner. The money they provided would double in value before being delivered to their partner, while the remaining amount was theirs to keep. Their partner would simultaneously make the same decision, and the participant would receive twice the amount provided by the partner. Depending on the experimental condition assigned, participants were either informed, or not informed, of their partner’s group membership. Participants were not informed of how much they had earned at the end of each transaction. Their decisions (how much they gave to their partner in each round) would not be disclosed to other participants or to the experimenter. In each trial, participants decided both how much they would contribute to their partner (“reward”) and were asked how much they expected their partner to contribute to them (“expectation”).

To ascertain whether behavior exhibited by group members is cooperation or derogation, it was necessary to include a control condition (Balliet et al., 2014; Schiller, Baumgartner, & Knoch, 2014). In Jin and Yamagishi’s (1997) paradigm, in the control condition, participants were not informed about their partner’s group belongingness. The cooperation in this condition reflected the baseline of participants’ cooperative tendency (Jin & Yamagishi, 1997). In this paper, this control condition is called as the “unknown” condition.

Manipulating participants’ knowledge of their partner’s group membership can reveal the motivations underlying cooperation and derogation. Specifically, if both participants know that they are partnered with an ingroup member in the PD game (i.e., the bilateral condition), the resulting cooperative behavior is likely to be motivated by the expectation of reciprocity (Yamagishi, Mifune, Liu, & Pauling, 2008). However, when only one participant is informed of their partner’s group membership, but is told that their partner is unaware of which group the participant belongs to (i.e., the unilateral condition), any cooperative behavior is motivated by factors other than the expectation of reciprocity. Yamagishi, Jin, and Kiyonari (1999) argued that because the salience of group boundaries allows participants to expect generalized exchanges within their group, ingroup cooperation can be motivated by reciprocity, defined as the expectation that ingroup members might return their contribution to (e.g., Yamagishi & Kiyonari, 2000; Kiyonari, 2002). The asymmetry of information related to a partner’s group membership prevents expectations of reciprocal return, because the participants know that their partner does not have knowledge of the participant’s own group belonging. In addition, group boundaries put limitations on the exchange. Therefore, reciprocity cannot be expected when participants are informed about their partner belonging to the outgroup. Derogation or non-cooperative behavior in situations where participants are bilaterally aware of their group membership to an outgroup is regarded as reciprocal defection, which is motivated by the fear of being defected upon by an outgroup member. In contrast, derogation in situations where group membership in the outgroup is unilaterally informed (only the participant knows their partner’s group membership) reflects active aggression toward that outgroup. Thus, using Jin and Yamagishi’s (1997) paradigm, we can clarify which types of cooperation and derogation are triggered by cues indicating an outgroup threat.

After reading the instructions, participants played six trials of the PD game. The total number of trials to be played was not shared with the participants in advance, and the order of the trials was randomized. In each trial, participants’ knowledge of their transaction partner’s group membership was manipulated to assess expectations of reciprocity. In the unknown condition participants were not informed of the group membership of their partner. In two trials, the group membership (in/outgroup) of both players was simultaneously revealed to participants and their partners. In this bilateral condition, expectations of reciprocity would likely be enhanced only toward ingroup members (Jin & Yamagishi, 1997; Kiyonari, 2002). In the remaining two trials, called the unilateral condition, participants played with a member of their own or another group who were not informed of the participants’ group belonging.

When all trials were completed, participants received the post-experiment questionnaire booklet including demographic questions. Once participants had completed all questionnaires, they signed a receipt to collect their earnings from the study and left the laboratory. The entire experiment took appropriately one hour to complete.

References

Balliet, D., Wu, J., De Dreu, Carsten K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*, 1556-1581. doi: 10.1037/a0037737

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/BF03193146

Jin, N., & Yamagishi, T. (1997). Group heuristics in social dilemma. *Japanese Journal of Social Psychology*, *12*, 190–198 (in Japanese with an English abstract).

Kiyonari, T. (2002). Expectations of a generalized exchange system and ingroup favoritism: An experimental study of bounded reciprocity. *The Japanese Journal of Psychology*, *73*, 1-9. (in Japanese) doi: 10.4992/jjpsy.73.1

Schiller, B., Baumgartner, T., & Knoch, D. (2014). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*, *35*, 169–175. doi: 10.1016/j.evolhumbehav.2013.12.006

Sugiura, H., Mifune, N., Tsuboi, S., & Yokota, K. (2017). Gender differences in intergroup conflict: The effect of outgroup threat priming on social dominance orientation. *Personality and individual differences*, *104*, 262–265. doi: 10.1016/j.paid.2016.08.013

Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity: Ingroup boasting and ingroup favoritism. *Advances in Group Processes*, *16*, 161–197.

Yamagishi, T., & Mifune, N. (2009). Social exchange and solidarity: in-group love or out-group hate? *Evolution and Human Behavior,* *30*, 229-237. doi: 10.1016/j.evolhumbehav.2009.02.004