

# Psychological mechanisms underlying ingroup favouritism in cooperation: Revisiting the reputation management and expectation hypotheses

*Group Processes & Intergroup Relations*

1–17

© The Author(s) 2024



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/13684302241239860  
journals.sagepub.com/home/gpi



HirotaKa Imada,<sup>1,2</sup>  Nobuhiro Mifune<sup>2</sup>  
and Hiroshi Shimizu<sup>3</sup>

## Abstract

According to the theory of bounded generalized reciprocity (BGR), intergroup contexts afford individuals the assumption that indirect reciprocity is bounded by group membership, and this shapes ingroup favouritism in cooperation. The assumption of bounded indirect reciprocity is hypothesized to result in ingroup favouritisms via two pathways: it leads people to behave in ways that earn and maintain a positive reputation in the eyes of ingroup, but not outgroup, members (the reputation management hypothesis), and it leads individuals to expect other ingroup members to be more cooperative than outgroup members (the expectation hypothesis). In other words, BGR offers two parallel psychological explanations for why people display ingroup favouritism. While the latter hypothesis has gained much experimental support, evidence for the former is rather scarce. Here, we report a direct test of both the reputation management hypothesis and the expectation hypothesis using two economic games. Overall, we found support for the expectation hypothesis, but not for the reputation management hypothesis.

## Keywords

bounded generalized reciprocity, indirect reciprocity, ingroup favouritism, intergroup cooperation, reputation

Paper received 24 June 2023; revised version accepted 27 February 2024.

Cooperation is the key to solving many societal and global challenges, such as climate change, social inequality, and intergroup conflicts. Nevertheless, cooperation is hard to achieve (e.g., Hardin, 1968). One of the critical factors that make it hard is that those who must cooperate often come from different groups. Previous

<sup>1</sup>Royal Holloway, University of London, UK

<sup>2</sup>Kochi University of Technology, Japan

<sup>3</sup>Kwansei Gakuin University, Japan

## Corresponding author:

HirotaKa Imada, Department of Psychology, Royal Holloway, University of London, Egham, TW20 0EX, UK.  
Email: himada2022@gmail.com

studies have consistently documented ingroup favouritism (Balliet et al., 2014; Yamagishi et al., 1999), the tendency to be more cooperative toward ingroup members than toward outgroup members, and researchers have long investigated the psychological mechanisms underlying this tendency (for reviews, see Balliet et al., 2014; Everett et al., 2015a). To better understand the psychological underpinning of ingroup favouritism, we revisit the two parallel hypotheses derived from the bounded generalized reciprocity perspective (Yamagishi et al., 1999): the reputation management hypothesis (Mifune et al., 2010; Yamagishi & Mifune, 2008) and the expectation hypothesis (Yamagishi et al., 1999). In the present research, we report an investigation of how reputational concern and expected cooperation from others together shape ingroup favouring tendencies.

Utilizing economic games such as the prisoner's dilemma, previous studies have documented ingroup favouritism in a wide range of intergroup relationships, from arbitrarily created experimental intergroup contexts (e.g., the minimal group paradigm; Otten, 2016; Tajfel et al., 1971) to natural intergroup situations in, for instance, political (Koopmans & Rebers, 2009) and national (Romano et al., 2021) contexts. Moreover, Romano et al. (2021) observed ingroup favouritism across 42 different countries, and Lazić et al.'s (2021) meta-analysis has found that children and adolescents also display the tendency in economic games. These findings point to the ubiquity of ingroup favouritism.

However, previous studies have shown that people do not always favour ingroup members over outgroup members. More specifically, people do not demonstrate ingroup favouritism when their group membership is not known to all parties (for reviews, see Balliet et al., 2014; Yamagishi et al., 1999). That is, people are more cooperative towards ingroup members than towards outgroup members only when they and their ingroup partner(s) both know that they belong to the same group. Yamagishi et al. offered an evolutionary and social psychological explanation for the conditional nature of ingroup

favouritism: the bounded generalized reciprocity perspective (BGR; Yamagishi et al., 1999).

According to BGR, intergroup contexts trigger the group heuristic that ingroup members belong to the same system of indirect reciprocity and, correspondingly, a favour given to an ingroup member is likely to be reciprocated by another member of the group. Individuals who belong to a system of indirect reciprocity base their cooperation on others' reputation (Milinski et al., 2002; Nowak & Sigmund, 2005); they cooperate with those who have a good reputation and withhold cooperation from those who have a bad reputation. In such a system, it could be fatal to establish a bad reputation, as this leads to negative reputational consequences, such as ostracism. As such, it is more costly to erroneously assume that one's noncooperation will not be detected when it can be than to erroneously assume that one's noncooperation will be detected when it cannot be. Thus, based on an error-management analysis (Haselton & Buss, 2000), the group heuristic provides an adaptive advantage by discouraging people from making more costly mistakes; when people know that their shared ingroup membership is known to their interaction partners, they display cooperation so that they can remain in and benefit from the system of indirect reciprocity.

Based on BGR, previous studies have examined the two proximate psychological mechanisms of ingroup favouritism: expected cooperation from an immediate interaction partner (Yamagishi et al., 1999) and reputational concern (Mifune et al., 2010; Yamagishi & Mifune, 2008). Yamagishi et al. (1999) hypothesized that people cooperate more with ingroup members than with outgroup members because they can expect more cooperation from ingroup members (i.e., those who are part of the same system of indirect reciprocity) than from outgroup members (the expectation hypothesis<sup>1</sup>). They tested this hypothesis using a group membership knowledge manipulation in economic games. This manipulation consists of the following two treatment conditions. In the common knowledge condition, participants are told the group membership of their interaction

partner, and they are also told that their own group membership is known to their partner. By contrast, in the unilateral knowledge condition, although participants are told the group membership of their partner, they are told that their own group membership is unknown to their partner. In this latter condition, participants are completely anonymous, and so cannot expect ingroup partners to be cooperative based on their shared group membership. Thus, BGR predicts that ingroup favouritism driven by expected cooperation will emerge in the common but not in the unilateral knowledge condition. Several previous experimental studies have supported this prediction (e.g., Jin et al., 1996; Yamagishi & Kiyonari, 2000; Yamagishi et al., 1999). In addition, previous studies have also reported that people do indeed expect ingroup members to be more cooperative in economic games than outgroup members, but this holds only in the common knowledge condition, not in the unilateral knowledge condition (e.g., Jin et al., 1996; Yamagishi et al., 1999). Thus, expected cooperation has been shown to be a key driver of ingroup favouritism in the common knowledge condition.

However, Simpson (2006) documented ingroup favouritism that could not be accounted for by the expectation hypothesis. In the basic prisoner's dilemma game, two individuals are paired, and they each receive the same amount of money from an experimenter. Their task is to decide how much of that money they will send to their partner, with the understanding that the money sent will be doubled before it is received by their partner. Importantly, in the basic game, the two individuals make their decisions simultaneously, and so do not know in advance how much their partner will transfer to them. Simpson (2006) modified the game such that the two players made their decisions sequentially. In this situation, based on the expectation hypothesis, the second player should not display ingroup favouritism when they already know that their partner has cooperated with them, even under a common knowledge treatment. Further, this should be true whether the partner is a member of the ingroup or the outgroup. However, contrary to the

hypothesis, Simpson (2006) showed that second players were still more cooperative with ingroup partners than outgroup partners. This finding suggests that expected cooperation cannot be the sole driver of ingroup favouritism, and it calls for an explanation for why second movers in a sequential prisoner's dilemma game should favour ingroup members over outgroup members.

To account for these findings, Yamagishi and Mifune (2008) proposed the reputation management hypothesis<sup>2</sup> as an alternative BGR-based proximate explanation for ingroup favouritism. They pointed to the crucial role of reputational concern in the system of indirect reciprocity, arguing that people are motivated to maintain a positive reputation in order to remain in the group and benefit from that system (Mifune et al., 2010; Yamagishi & Mifune, 2008). In other words, the reputation management hypothesis focuses on reputational concern, rather than expected cooperation, as the core psychological underpinning of ingroup favouritism.

Based on the reputation management hypothesis, Mifune and Yamagishi (2008) predicted that people experience increased reputational concern when interacting with ingroup members, and this leads to increased ingroup cooperation. As such, the reputation management hypothesis allowed them to explain ingroup favouritism when expected cooperation is structurally ruled out (e.g., ingroup favouritism among the second players in the sequential prisoner's dilemma game, as in Simpson, 2006). They tested the hypothesis using a dictator game in which participants had a chance to benefit their partner, but their partner did not have a chance to benefit them (i.e., expected cooperation was structurally removed). Their experiment supported the prediction: ingroup favouritism in the dictator game emerged in a common knowledge condition but not in a unilateral knowledge condition. Moreover, Mifune et al. (2010) had participants play the dictator game either with an ingroup or an outgroup member under unilateral knowledge conditions, with the presence versus absence of "stylized eyes" as a manipulation of a monitoring cue. More specifically, participants were briefly

presented with an image of watching eyes of a kabuki actor (i.e., eyes with kumadori make-up that emphasizes muscles and veins around eyes) or an image of the ancient Egyptian Eye of Horus. These watching eye images were designed to trigger reputational concern. Two experiments yielded converging evidence that ingroup favouritism occurred only when the cue was present, supporting the reputation management hypothesis. However, these studies did not directly measure reputational concern, so it remains unclear whether it is actually reputational concern that was responsible for the emergence of ingroup favouritism. Finally, two correlational studies also provide anecdotal evidence in support of the reputation management hypothesis, showing that trait reputational concern is correlated with ingroup favouritism when group membership is common knowledge (Kajiwara et al., 2022; Mifune & Yamagishi, 2015). Overall, then, previous studies suggest that reputational concern should drive ingroup favouritism—at least when expected cooperation is ruled out. Yet, the current literature lacks direct evidence of the mediational role played by reputational concerns as predicted by BGR, that is, a common knowledge manipulation that increases ingroup but not outgroup cooperation via increased reputational concern.

In summary, based on BGR, there are two parallel hypotheses regarding the proximate psychological mechanism driving ingroup favouritism: expected cooperation and reputational concern. Yet, direct evidence of the role played by reputation concerns is lacking, and the relative importance of the two psychological mechanisms has been understudied. To address these issues, we conducted two parallel experiments. In Study 1, we used a giving game, where participants had a chance to unidirectionally cooperate with another person (expected cooperation was structurally ruled), and we explored whether or not reputational concern mediated the relationship between our knowledge manipulation and ingroup favouritism. In Study 2, we employed a prisoner's dilemma game in which the common knowledge manipulation could lead to ingroup

favouritism via increased expected cooperation, reputational concern, or both. Here, our goal was to examine the relative importance of these two mechanisms.

Note that while the BGR perspective is based on the idea that people intuitively assume that indirect reciprocity is bounded by group membership, Romano and colleagues argued that indirect reciprocity is not always assumed to be bounded by group membership (Romano et al., 2017). Romano et al. (2017) used an experimental paradigm in which participants first play a cooperation game (the dictator game or the prisoner's dilemma game), and their decision in the game would be communicated to a future interaction partner in the next game. With this experimental paradigm, they demonstrated that reputational concern promoted cooperation in the first game, regardless of the group membership of the interaction partner. This experimental finding was inconsistent with the BGR perspective and suggests that people perceived indirect reciprocity to be unbounded by group membership. Nevertheless, Imada et al. (2023) reconciled the conflict by proposing the dynamic indirect reciprocity perspective, arguing that (a) by default, indirect reciprocity is assumed to be bounded by group membership when group membership is the only information on which they can base their cooperation decision, but (b) this default assumption can be overridden (as originally demonstrated by Yamagishi & Kiyonari, 2000)—with indirect reciprocity assumed to be unbounded—when there are additional cues suggesting the potential for future interactions and benefits. In the present research, we aimed to further elucidate the proximate mechanism(s) of ingroup favouritism predicted by BGR, and thus we focused on situations in which group membership was the sole cue.

Also note that the social identity perspective (Tajfel & Turner, 1979) and the BGR perspective have sparked a lively debate about the psychological mechanism behind ingroup favouritism in economic games (Balliet et al., 2014; Everett et al., 2015a; Simpson, 2006; Yamagishi et al., 1999). While there are some experiments supporting the

former (e.g., Hackel et al., 2017), the large-scale meta-analysis by Balliet et al. (2014) yielded evidence in favour of BGR, and there is abundant evidence demonstrating a lack of relationship between social identification and ingroup favouritism in cooperation (Imada et al., 2023; Romano et al., 2017; Yamagishi & Kiyonari, 2000; Yamagishi & Mifune, 2008). In light of this, our focus was on the BGR perspective, and we did not intend to compare it against the social identity perspective in the present research. That said, we did measure social identification, and we reported analyses relevant to the social identity perspective in the online supplemental results. Overall, across the two studies, the strength of social identification was not associated with cooperation.

## Study 1

In the present study, we aimed to extend the work of Yamagishi and Mifune (2008) and directly test the reputation management hypothesis. Thus, we had two hypotheses:

H1a: Individuals will cooperate more with an ingroup partner than with an outgroup partner in the common knowledge condition.

H1b: Individuals will cooperate no more or less with an ingroup partner than with an outgroup partner in the unilateral knowledge condition.

Further, and going beyond previous studies supporting the reputation management hypothesis, we also measured the extent to which participants felt reputational concern while making a cooperation decision, and we had the following hypotheses regarding reputational concern:

H2a: Individuals will experience more reputational concern when playing with an ingroup member than with an outgroup member in the common knowledge condition.

H2b: Individuals will experience no more or less reputational concern when playing

with an ingroup member than with an outgroup member in the unilateral knowledge condition.

In addition, we tested the moderated mediation effect of reputational concern on the relationship between knowledge manipulation and cooperation, using a multilevel path analysis. We preregistered the hypotheses, analytic plans, study procedures, and target sample size at the Open Science Framework (<https://osf.io/d6mze>; study material, data, analysis code, and supplemental results are stored at <https://osf.io/kq3x2/>).

## Methods

The present study employed a 2 (group: ingroup vs. outgroup)  $\times$  2 (knowledge: common knowledge vs. unilateral knowledge) within-subjects design. We conducted an a priori power analysis with the “Superpower” R package (Lakens & Caldwell, 2021), using means and standard deviations from a previous study with a similar design (Imada et al., 2023),<sup>3</sup> and it revealed that 130 participants would be sufficient to detect an interaction effect with 90% statistical power and  $\alpha = 5\%$ . In order to ensure statistical power after data exclusion, we recruited 150 undergraduate students from a British university in exchange for partial course credit. Four participants did not finish the study, thus we had 146 complete responses. After applying the preregistered completion-time-based data exclusion criterion, we were left with 145 participants ( $M_{\text{age}} = 19.48$ ,  $SD_{\text{age}} = 5.01$ ; 17 males, 128 females).

After giving consent, participants took part in an online survey, which was advertised as being about artistic preferences and decision-making. They were told that the survey consisted of two parts: an artistic preference task and a decision-making task. Following the procedure employed by Everett et al. (2015b), participants were first presented with several pairs of paintings and asked to choose the one they preferred. After they completed the task, they were given bogus feedback that there were two groups of people varying in their artistic preferences, and that they

were categorized into Group A. Participants read a brief description of their group and, in order to strengthen the minimal group induction, were asked to provide an example in which their behaviour seemed to be consistent with the description. They then answered six questions measuring social identification with the minimal group. For this purpose, we employed the six-item scale used in Leonardelli and Brewer (2001), where participants were asked to respond on a 6-point Likert-type scale (1 = *strongly disagree*, 6 = *strongly agree*). A sample item is, "I feel that Group A is an important reflection of who I am" ( $\alpha = .67$ ).

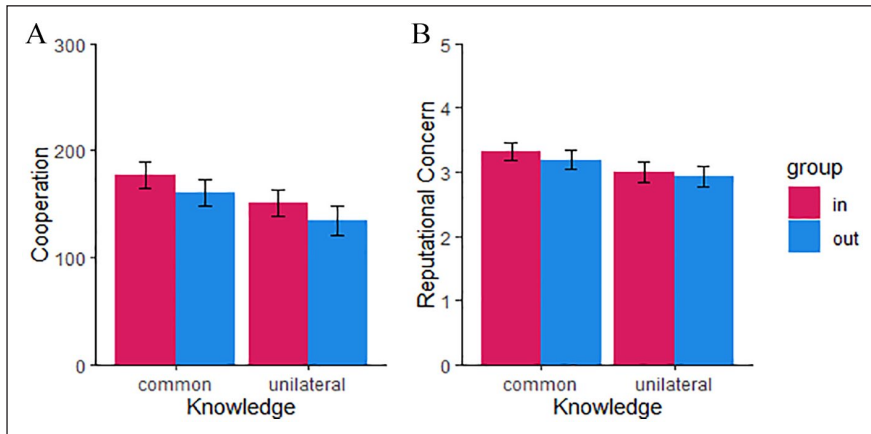
Next, participants read the instructions about the giving game. In the game, participants were given 300 pence and paired with another person. Their task was to decide how much of the 300 pence they would like to transfer to their partner, knowing that each pence they decided to send would be doubled before it was given to the partner. They were instructed that they would complete the game several times, each with a different person. To incentivize them, we told participants that one participant would be randomly selected to receive the actual payment of the money they earned in the study (i.e., the amount of money they decided to keep for themselves). Note, however, that although participants were told they would be matched with others online, they in fact were not matched with anyone during the study.

Participants played the game four times, and we orthogonally manipulated the group membership of the partner and whether or not the participants' reputation was at stake. To accomplish the ingroup/outgroup manipulation, participants were instructed either that their partner was from the same group (Group A) or from the other group (Group B), respectively. In the unilateral knowledge condition, they were told that while they knew the group membership of their partner, their partner did not know anything about them, including their group membership. In other words, participants were completely anonymous, so their reputation was not at stake in that condition. In the common knowledge

condition, by contrast, participants were told that they and their partner both knew about the group membership of each other. Thus, in this condition, their reputation was at stake. Each participant completed four games, with a different combination of partner group membership and knowledge condition in each game. Different participants encountered the games in different random orders. After making each decision, they were asked to complete a short questionnaire measuring reputational concern (Wu et al., 2015). In addition, for exploratory purposes, we asked participants how much they believed their partner expected them to transfer, and this is not considered further here. To measure reputational concern, we used four questions from the reputational concern scale (Wu et al., 2015; e.g., "I did not consider what my partner would say about me"; all  $\alpha > .79$ ). The scale was measured by a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). Since we asked the same questions four times, we inserted three filler questions (e.g., "I felt happy during the task") between the reputational concern items. After participants completed the game four times, they provided demographic information (e.g., gender, age, nationality, and first language) and were fully debriefed.

## Results and Discussion

We first conducted a 2 (group: ingroup vs. outgroup)  $\times$  2 (knowledge: common knowledge vs. unilateral knowledge) within-subjects ANOVA on cooperation (Figure 1A). The two main effects were both significant,  $F_s > 15.69$ ,  $p_s < .001$ , all  $\eta_{ps}^2 > .10$ . As expected, the analysis yielded a significant interaction,  $F(1, 144) = 4.01$ ,  $p = .047$ ,  $\eta_p^2 = .05$ . Following the preregistration, we examined the simple main effect of group in each of the knowledge conditions. We tested H1a and H1b with simple main effect analyses, using the cutoff point of  $\alpha = .025$ , because in order for BGR and the reputation management hypothesis to be supported, two hypotheses (H1a and H1b) should be supported. We applied this cutoff in Study 2 as well. In the common

**Figure 1.** (A) Cooperation and (B) reputational concern by condition.

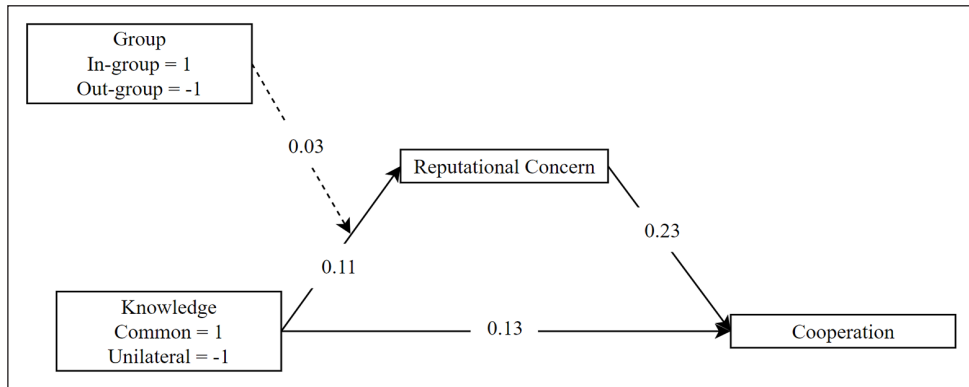
Note. Error bars indicate 95% confidence intervals.

knowledge condition, participants were more cooperative towards an ingroup member ( $M = 142.34$ ,  $SD = 62.41$ ) than towards an outgroup member ( $M = 120.69$ ,  $SD = 57.62$ ),  $F(1, 144) = 25.87$ ,  $p < .001$ ,  $\eta_p^2 = .15$ . In the unilateral knowledge condition, they were also more cooperative toward an ingroup member ( $M = 126.26$ ,  $SD = 57.16$ ) than toward an outgroup member ( $M = 114.09$ ,  $SD = 58.06$ ),  $F(1, 144) = 10.09$ ,  $p = .002$ ,  $\eta_p^2 = .07$ . These results suggest that the magnitude of ingroup favouritism was stronger in the common knowledge condition than in the unilateral knowledge condition, but that ingroup favouritism was still present in the latter condition. These results supported H1a but not H1b. Since H1b was a null hypothesis, we also conducted a Bayesian ANOVA with the Jeffreys prior, and found that it strongly favoured the alternative hypothesis (i.e., ingroup and outgroup cooperation are different),  $BF_{10} = 13.57$ .

We then carried out a 2 (group: ingroup vs. outgroup)  $\times$  2 (knowledge: common knowledge vs. unilateral knowledge) within-subjects ANOVA on reputational concern (Figure 1B). The main effects of group and knowledge were both significant,  $F_s > 6.53$ ,  $p_s < .011$ ,  $\eta_{ps}^2 > .04$ . However, the interaction effect was not significant,  $F(1, 144) = 0.72$ ,  $p = .400$ ,  $\eta_p^2 = .004$ . Thus, H2a and H2b were not supported, inconsistent with the reputation management hypothesis.

We conducted an exploratory multilevel moderated mediation analysis (see Figure 2) using Mplus Version 7.3 (Muthén & Muthén, 2017). We used maximum likelihood parameter estimates with standard errors that are robust to nonnormality and nonindependence of observations (MLR estimator). We report within-subjects level path coefficients and full results in our online supplemental results (intraclass correlations for cooperation and reputational concern = .72 and .64, respectively). Note that since we manipulated the exogenous variable (i.e., knowledge), there should not be any confounding variables that influence our estimation of the causal relationship between the knowledge manipulation and the mediator. Yet, we must assume that there are no confounding variables between the mediator and the endogenous variable (i.e., cooperation). The effect of knowledge on reputational concern was significant,  $\beta = .11$ , 95% CI [0.04, 0.18],  $p = .003$ . In addition, the effect of reputational concern on cooperation was significant,  $\beta = .23$ , 95% CI [0.10, 0.37],  $p = .001$ . However, the path between the knowledge manipulation and reputational concern was not moderated by the group manipulation,  $\beta = .03$ , 95% CI [-0.04, 0.10],  $p = .400$ . The direct effect of knowledge on cooperation was significant,  $\beta = .13$ , 95% CI [0.06, 0.20],  $p < .001$ . The overall mediation effect was significant,  $\beta = .03$ , 95% CI [0.001, 0.05],  $p = .032$ .

**Figure 2.** Mediation diagram and standardized path coefficients: Study 1.



Note. Solid and dotted lines indicate significant and nonsignificant paths, respectively.

Given the significant direct effect, we thus found a partial mediation effect of reputational concern. The results together suggest that the group membership knowledge manipulation increases cooperation in part via increased reputational concern, regardless of group membership of the interaction partner. This does not support the reputation management hypothesis.

In Study 1, we found that individuals displayed ingroup favouritism not only in the common knowledge condition but also in the unilateral knowledge condition. However, the degree of ingroup favouritism was stronger in the former condition. This is not in line with BGR, as it strictly predicts that the emergence of ingroup favouritism is conditional on common knowledge of the shared ingroup membership. We turn to this issue in the General Discussion section. In addition, we did not find support for the reputation management hypothesis, revealing that participants did not experience a higher level of reputational concern when playing with an ingroup partner than with an outgroup partner in the common knowledge condition.

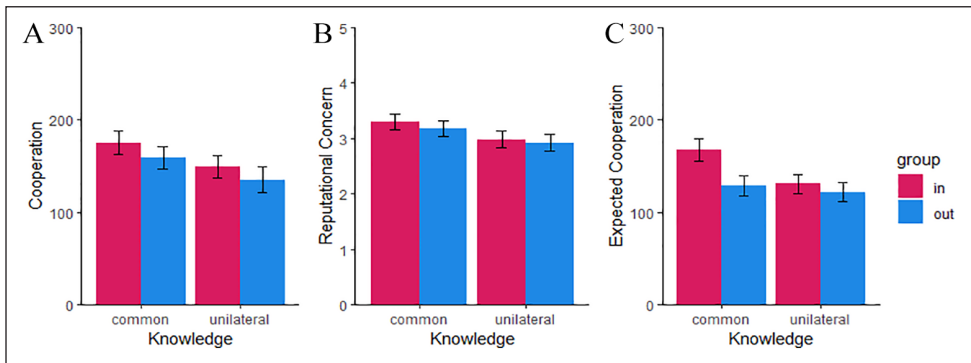
The multilevel moderated mediation analysis was conducted to explore the processes by which knowledge manipulation and reputational concern shaped intergroup cooperation. Contrary to the reputation management hypothesis, we found

that common knowledge manipulation increased both ingroup and outgroup cooperation via increased reputational concern. Taken together, Study 1 suggests that the reputation management hypothesis does not explain ingroup favouritism in cooperation.

## Study 2

Study 1 used a giving game in which participants could not expect to receive any cooperation from their partner. Thus, the other important psychological mechanism (besides reputational concerns) presumed to underlie ingroup favouritism—expected cooperation—was structurally removed. In the present study, by contrast, we specifically explored the relative influence of reputation and expected cooperation on ingroup favouritism using a prisoner's dilemma game in which two individuals simultaneously decide the extent to which they will cooperate with each other. In the prisoner's dilemma game, the common knowledge treatment can potentially lead to ingroup favouritism via both reputational concern and expected cooperation, thus simultaneously testing the two parallel hypotheses. Thus, in addition to the four hypotheses tested in Study 1, we also tested the following two hypotheses concerning expected cooperation:



**Figure 3.** (A) Cooperation, (B) reputational concern, and (C) expected cooperation by condition.

Note. Error bars indicate 95% confidence intervals.

H3a: Individuals will expect an ingroup partner to be more cooperative than an outgroup partner in the common knowledge condition.

H3b: Individuals will expect an ingroup partner to be no more or less cooperative than an outgroup partner in the unilateral knowledge condition.

The present study was identical to Study 1 except for the type of game; and our hypotheses and preregistration for Study 2 mirrored those for Study 1. We did not preregister H3a and H3b.<sup>4</sup>

We preregistered the hypotheses, analytic plans, study procedures, and target sample size at the OSF (<https://osf.io/zpf4s>; we have study material, data, analysis code, and supplementary results stored at <https://osf.io/kq3x2/>).

## Methods

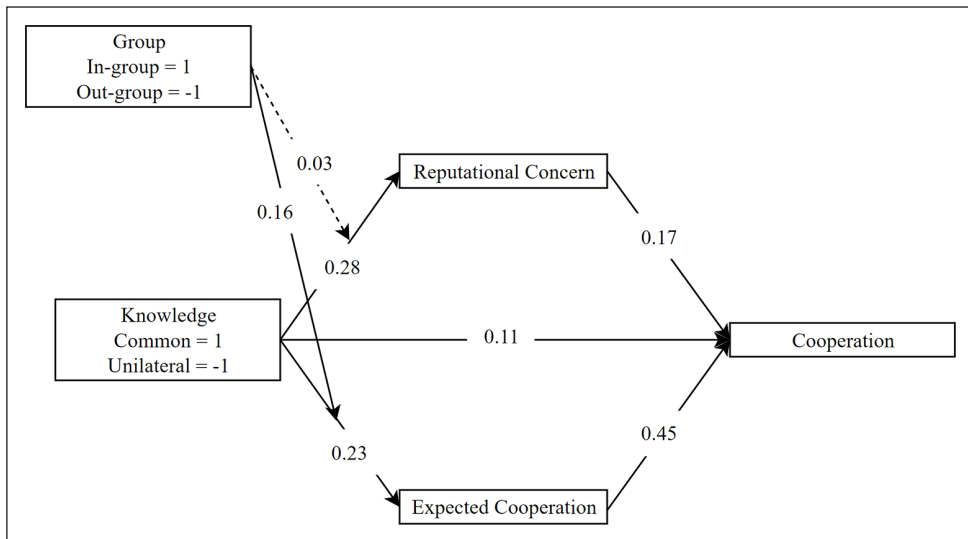
The procedure of the present study was identical to that of Study 1, except that we used a prisoner's dilemma game instead of the giving game. We determined the target sample size using the power analysis reported in Study 1. Given we had some dropouts and excluded data in Study 1, we again overrecruited 150 undergraduate students from a British university in exchange for partial course credit; 143 participants finished the entire study. After applying preregistered data exclusion criteria based on

study completion time, we had 142 completed responses ( $M_{\text{age}} = 19.56$ ,  $SD_{\text{age}} = 5.50$ ; 22 males, 119 females).

The rules of the prisoner's dilemma game were as follows: participants were paired with another person, and they were each endowed with 300 pence. They were then asked to decide how much money they would like to transfer to their partner, knowing that each pence they transferred would be doubled before it was given to their partner. The difference between the giving game (Study 1) and the prisoner's dilemma game (Study 2) is that in the latter, the participant's partner had the opportunity to make a similar decision about transferring money to the participant. We measured expected cooperation by directly asking participants how much they believed their partner had transferred to them. As in Study 1, we measured social identification (Leonardelli & Brewer, 2001) and reputational concern (Wu et al., 2015), and the scales both had satisfactory reliability (all  $\alpha > .72$ ).

## Results and Discussion

We conducted a 2 (group: ingroup vs. outgroup)  $\times$  2 (knowledge: common knowledge vs. unilateral knowledge) within-subjects ANOVA on cooperation (Figure 3A). The main effects of group and knowledge were both significant,  $F_s > 16.04$ ,  $p_s < .001$ ,  $\eta_{ps}^2 > .10$ . However, the interaction effect was not significant,  $F(1,$

**Figure 4.** Mediation diagram and standardized path coefficients: Study 2.

Note. Solid and dotted lines indicate significant and nonsignificant paths, respectively.

139) = 0.003,  $p = .957$ ,  $\eta_p^2 < .001$ . H1a and H2b were not supported. The results thus suggest that ingroup favouritism was not dependent on the group membership knowledge manipulation, inconsistent with BGR.

We then conducted a 2 (group: ingroup vs. outgroup)  $\times$  2 (knowledge: common knowledge vs. unilateral knowledge) within-subjects ANOVA on reputational concern (Figure 3B). Consistent with the analysis of cooperation, we found two significant main effects,  $F_s > 6.15$ ,  $p_s < .014$ ,  $\eta_{ps}^2 > .04$ , but the interaction term was not significant,  $F(1, 141) = 0.71$ ,  $p = .400$ ,  $\eta_p^2 = .005$ . Thus, H2a and H2b were not supported. These results did not support the reputation management hypothesis.

We also conducted the same analyses on expected cooperation. The 2  $\times$  2 within-subjects ANOVA revealed a significant interaction effect,  $F(1, 141) = 21.64$ ,  $p < .001$ ,  $\eta_p^2 = .13$ . The two main effects were also significant,  $F_s > 31.32$ ,  $p_s < .001$ ,  $\eta_{ps}^2 > .18$ . The simple main effect analyses showed that participants in the common knowledge condition expected an ingroup partner ( $M = 167.99$ ,  $SD = 73.76$ ) to be more cooperative than an outgroup partner ( $M = 128.33$ ,  $SD = 67.72$ ),  $F(1,$

141) = 46.06,  $p < .001$ ,  $\eta_p^2 = .25$ . In the unilateral knowledge condition, they did not expect an ingroup partner ( $M = 130.73$ ,  $SD = 61.51$ ) to be more cooperative than an outgroup partner ( $M = 122.00$ ,  $SD = 60.55$ ),  $F(1, 141) = 4.44$ ,  $p = .037$  ( $> .025$ ),  $\eta_p^2 = .03$ . These results supported both H3a and H3b. Since H3b was a null hypothesis, we conducted a Bayesian ANOVA and found inconclusive evidence as to whether the null hypothesis (H3b) or an alternative hypothesis (expected cooperation is different between the ingroup and outgroup conditions) would be more plausible,  $BF_{10} = 1.05$  (not preregistered). Thus, while our preregistered test supported the expectation hypothesis, the results should be taken cautiously; we did not have strong evidence that expected cooperation from an ingroup or from an outgroup partner did not substantially differ in the unilateral knowledge condition.

Next, we conducted a multilevel mediation analysis (see Figure 4). Intraclass correlations for cooperation, reputational concern, and expected cooperation were as follows: ICC = .61, .67, and .52, respectively. Note that, as in Study 1, we assume that there are no confounding variables between the mediators and the endogenous

variable (i.e., cooperation). Additionally, we assumed that there was no causal relationship between the two mediators.<sup>5</sup>

The effect of knowledge on reputational concern was significant,  $\beta = .28$ , 95% CI [0.19, 0.36],  $p < .001$ . In addition, the effect of reputational concern on cooperation was significant,  $\beta = .17$ , 95% CI [0.07, 0.27],  $p = .001$ . However, the path between the knowledge manipulation and reputational concern was not moderated by the group manipulation,  $\beta = .03$ , 95% CI [-0.04, 0.03],  $p = .398$ . Yet, the direct effect of knowledge on cooperation was significant,  $\beta = .11$ , 95% CI [0.06, 0.20],  $p = .002$ . The overall mediation effect was significant,  $\beta = .05$ , 95% CI [0.01, 0.08],  $p = .005$ . Given the significant direct effect, we thus found a partial mediation effect of reputational concern.

The effect of knowledge on expected cooperation was significant,  $\beta = .23$ , 95% CI [0.16, 0.30],  $p < .001$ . In addition, the effect of expected cooperation on cooperation was significant,  $\beta = .45$ , 95% CI [0.34, 0.57],  $p < .001$ . Furthermore, the path between the knowledge manipulation and expected cooperation was moderated by the group manipulation,  $\beta = .16$ , 95% CI [0.10, 0.23],  $p < .001$ . The overall mediation effect was significant,  $\beta = .10$ , 95% CI [0.06, 0.15],  $p < .001$ . Given the significant direct effect, we thus found a moderated partial mediation effect of expected cooperation.

In Study 2, ingroup favouritism was present regardless of group membership knowledge condition. This is inconsistent with BGR. In addition, we did not find support for H2a and H2b. The exploratory moderated mediation analysis also showed that the indirect effect of knowledge on cooperation via reputation concern was not influenced by group membership of the interaction partner. These results overall did not support the reputation management hypothesis. The expectation hypothesis obtained experimental support; people expected an ingroup member to be more cooperative than an outgroup member in the common knowledge condition, but not in the unilateral knowledge condition. Yet, as noted earlier, the effect of group on expected cooperation in the

unilateral knowledge condition warrants a cautious interpretation. Moreover, we found a moderated (partial) mediation effect via expected cooperation, in line with the expectation hypothesis. The results of the present study, therefore, suggest that the expectation hypothesis offers a more plausible psychological explanation for ingroup favouritism in prisoner's dilemma games than does the reputation management hypothesis.

## General Discussion

According to BGR, shared group membership leads people to believe that ingroup members, but not outgroup members, belong to the same system of indirect reciprocity. Yamagishi et al. (1999) argued that people have acquired this default assumption as it helps them avoid earning a negative reputation within their group by acting selfishly. This further promotes the assumption that people should maintain a positive reputation in the eyes of ingroup members (the reputation management hypothesis) and that other ingroup members will likely cooperate with them because they too should strive to maintain a positive reputation (the expectation hypothesis).

Previous studies have consistently documented ingroup favouritism under the common knowledge treatment (i.e., one's group membership is known to their interaction partner), and BGR offers two hypotheses as to what psychological processes underlie the phenomenon. Since the current literature lacks a direct examination of the reputation management hypothesis (Mifune et al., 2010; Mifune & Yamagishi, 2015; Yamagishi & Mifune, 2008), we aimed to test it in two games that varied in whether the game partner had the chance to benefit the participant or not. We found ingroup favouritism both in the common and the unilateral knowledge conditions, in both the giving game (Study 1) and the prisoner's dilemma game (Study 2), which is inconsistent with the BGR perspective. That is, our studies showed that the emergence of ingroup favouritism was not conditional on the group membership knowledge manipulation.

Furthermore, we did not find support for the reputation management hypothesis. More specifically, in both studies, we found that the common knowledge treatment increased reputational concern for both ingroup and outgroup partners. Contrastingly, the expectation hypothesis was supported in Study 2; participants expected an ingroup partner to be more cooperative than an outgroup partner in the common knowledge condition, but not in the unilateral knowledge condition. Below, we discuss the implications of our findings for the reputation management hypothesis, the expected cooperation hypothesis, and the bounded generalized reciprocity perspective.

### *The Reputation Management Hypothesis*

Regarding the reputation management hypothesis, past evidence supporting it was either correlational (Kajiwara et al., 2022; Mifune & Yamagishi, 2015) or did not directly measure reputational concern, and instead inferred its effect solely from the impact of experimental manipulations of relevant variables (Mifune et al., 2010; Yamagishi & Mifune, 2008). We aimed to overcome these limitations by offering direct experimental evidence for the hypothesis. In both Studies 1 and 2, we showed that participants felt a higher reputational concern in the common knowledge condition than in the unilateral knowledge condition, regardless of group membership of the interaction partner. In addition, we consistently found that the common knowledge treatment increased cooperation via reputational concern when interacting with both ingroup and outgroup partners. Overall, our studies suggest that the reputation management hypothesis does not explain the psychological mechanism underlying ingroup favouritism. Therefore, our studies call into question whether the belief that indirect reciprocity is bounded by group membership actually leads to ingroup favouritism through reputation management.

Yet, we argue that it would be premature to completely abandon the reputation management hypothesis. Previous studies in favour of the

hypothesis were conducted in Japan (Kajiwara et al., 2022; Mifune et al., 2010; Mifune & Yamagishi, 2015; Yamagishi & Mifune, 2008). Yamagishi et al. (1999) suggested that reputation management is particularly important in societies in which individual and intergroup mobility is limited, and Japan is characterized as such a society. Our participants, by contrast, were students at a British university, a majority of whom were from Western societies in which interpersonal and intergroup relationships are rather fluid (Thomson et al., 2018). This may explain why we did not find support for the reputation management hypothesis (but see Yamagishi et al., 2008). In other words, individuals in societies low in relational mobility may display both reputation- and expected-cooperation-driven ingroup favouritism.

In addition, according to the reputation management hypothesis, it is crucial for individuals not to earn a negative reputation in their group; the hypothesis is not directly concerned with reputations of positive valence. Since positive versus negative reputations are thought to be distinct (Wu et al., 2016), and have been separately manipulated (Imada, 2023; Imada, Hoptrow, & Abrams, 2021), when testing BGR's reputation management hypothesis, it seems sensible to operationally define and measure reputational concern as the extent to which people are concerned about earning a negative reputation. However, the items we used to measure reputational concern were phrased such that they appeared to be focused on one's positive reputation" (e.g., "It's important to me that my partner has a *positive evaluation* towards me," "I thought it's important that others will *accept* me"). This might have led to the reduced impact of the interaction between group membership and knowledge manipulation on reputational concern. We note that the existing correlational evidence for the reputation management hypothesis, in fact, focused on negative reputation, that is, fear of negative evaluation (Kajiwara et al., 2022; Mifune & Yamagishi, 2015). Overall, while our experiments did not yield evidence that reputation management concern predicts ingroup

favouritism, it is sensible to further examine the hypothesis with respect to potential boundary conditions (e.g., cultural differences and the valence of reputation).

### *The Expectation Hypothesis*

In Study 2, we tested the prediction derived from the expectation hypothesis: people favour ingroup members because they expect them to be more cooperative than outgroup members. Consistent with the hypothesis, in the common knowledge condition, expected cooperation was higher in the ingroup condition. In the unilateral knowledge condition, we did not find that people expected an ingroup partner to be more cooperative than an outgroup partner. Study 2 further offers valuable evidence in support of the expectation hypothesis in relation to the reputation management hypothesis. Previous experiments supporting the expectation hypothesis did not typically examine the effect of reputation management, and it was uncertain whether expected cooperation would explain ingroup favouritism while controlling for the effect of reputational concern. Our study closed this gap, finding experimental support for the expectation hypothesis while accounting for the effect of reputational concern.

### *Bounded Generalized Reciprocity*

In our studies, individuals displayed ingroup favouritism both in the common and unilateral knowledge conditions. They had to complete four games consecutively in a short period of time and they were constantly exposed to intergroup comparisons. This experimental setting may have bolstered the influence of ingroup-favouring social preference generally (e.g., social identification; Tajfel et al., 1971; Tajfel & Turner, 1979), and so contributed to shaping ingroup favouritism in the unilateral knowledge condition. In fact, Everett et al. (2015b) employed a similar procedure and found ingroup favouritism when one's decision was completely anonymous. They concluded that social preferences (i.e.,

ingroup biases not conditioned on situational cues such as the common knowledge manipulation) result in ingroup favouritism even under anonymity. The within-subjects design may have engendered psychological motivations or mechanisms that do not exist in one-shot games and, therefore, may not have been the best design for testing BGR.

As briefly noted in our introduction, the assumption of bounded indirect reciprocity is a default game strategy that people employ when group membership is the sole cue they have to rely on. Imada et al. (2023) have recently proposed the dynamic indirect reciprocity perspective, which posits that the perception of the realm of indirect reciprocity, that is, whether or not it is bounded by group membership, is dynamic and depends on available information. For instance, they experimentally demonstrated that when individuals are certain that cooperating with an ingroup or an outgroup member can bring future rewards, reputational concern increases both ingroup and outgroup cooperation (Imada et al., 2023; Romano et al., 2017). By contrast, when there was no potential for future rewards, individuals expected an immediate interaction with an ingroup partner to be more cooperative than an immediate interaction with an outgroup partner, and this led to ingroup favouritism. Our results are compatible with these findings. Consistent with Imada et al. (2023), our moderated mediation analyses suggested that the effect of reputational concern on cooperation was not bounded by group membership but that of expected cooperation was. Taken together, reputational concern may play a more important role when future rewards are implied, which is clearly not the case in a one-shot game situation. Rather, our results suggest that ingroup favouritism in a one-shot bidirectional game can be better explained by expected cooperation.

In previous literature, reputational concern and expected cooperation were thought to be independent of each other, and were pitted against one another as the potential psychological underpinning of prosocial behaviour (e.g., Imada, Hothrow, & Abrams, 2021; Wu et al., 2015).

More importantly, the reputation management hypothesis emerged as an explanation for ingroup favouritism among second movers in sequential prisoner's dilemma games, which could not be accounted for by the expectation hypothesis (Yamagishi & Mifune, 2008). Thus, it was assumed that reputational concern would function independently of expected cooperation. In addition, reputational concern and expected cooperation have been reported as being only weakly correlated with each other (Imada et al., 2023).<sup>6</sup> Drawing upon previous literature, we thus treated the two hypotheses as parallel and independent. However, one could argue that the assumption of group-bounded indirect reciprocity operates primarily by increasing reputational concerns among ingroup members, which in turn leads ingroup members to expect that other ingroup members will also be concerned about their reputations and so should behave cooperatively toward them. In other words, expected cooperation may be a derivative of reputational concern. The existing literature currently lacks a sound basis to support this argument, but the relationship between reputational concern and expected cooperation may be worth revisiting not only in intergroup but also in interpersonal cooperation contexts.

We would also like to note that because our aim was specifically to test BGR, we only discussed reputational concern and expected cooperation as the potential sources of ingroup favouritism in the common knowledge condition. However, there may be alternative explanations for why the common knowledge treatment leads to ingroup favouritism, for instance, group-specific social preferences (e.g., parochial altruism; Columbus et al., 2023; Yamagishi & Mifune, 2016). The elucidation of the psychological mechanisms of ingroup favouritism beyond BGR is one direction for future research.

In addition, while we focused on the minimal group context in order to examine the influence of group membership per se, naturally emerging groups substantially vary in, for example, stereotype content. Politically and morally conflicting intergroup contexts are often characterized as

involving severe polarization and hostile relations (Imada, Codd, & Liu, 2021; Waller & Anderson, 2021; Weisel & Böhm, 2015). In such contexts, prior investment in the ingroup and negative outgroup attitudes may contribute to ingroup favouritism. While Balliet et al. (2014) found that the strength of ingroup favouritism did not vary between minimal and natural groups, previous literature paid scant attention to whether the psychological mechanisms underlying ingroup favouritism would systematically vary depending on different intergroup contexts, suggesting a promising direction for future research.

### *Limitations*

Finally, we would like to note two methodological limitations. First, in both studies, we employed a rather weak incentive (i.e., just one participant in the study would receive the actual payment based on their behaviour in the game). This might have led to overestimated cooperation levels compared to those in more strongly incentivized studies. On the other hand, Romano et al. (2021) manipulated the presence (vs. absence) of incentives and found that the manipulation did not influence intergroup cooperation in a prisoner's dilemma game, which would seem to argue against the possibility that the weak incentives in the present research threatened the validity of the results. That being said, it would be sensible for future work to employ more strongly incentivized behavioural experiments.

Second, we did not directly manipulate the mediators and, for our multilevel moderated mediation analyses, we assumed that there were no confounding variables. As such, it would be sensible to reexamine the causal relationship between the mediators and cooperation by experimentally manipulating the mediators and testing their effect on cooperation (Ge, 2023). It would also be important to conduct causal mediation analyses in order to generate estimations of the causal effects that are robust against the violation of the assumption, and to quantify the robustness. Overall, it is of vital importance for future

research to seek robust estimations of the causal impacts of reputational concern and expected cooperation not only in intergroup but also in interpersonal contexts in general.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Hirota Imada  <https://orcid.org/0000-0003-3604-4155>

### Supplemental material

Data, material, analysis codes, and supplementary results associated with all studies reported in the article are available at the OSF (<https://osf.io/kq3x2/>).

### Notes

1. We would like to note two things. First, Yamagishi et al. (1999) never named their hypothesis concerning expectations. We have named it the “expectation hypothesis” simply to make it easy to contrast the two hypotheses. Second, Yamagishi et al. (1999) were not the first to point to the crucial role of expected cooperation in shaping ingroup favoritism—this dates back to several earlier studies (Karp et al., 1993; Rabbie & Horwitz, 1988; Rabbie et al., 1989). Yamagishi et al. (1999) were the first, however, to have offered the explanation for why people expect more cooperation from ingroup members than from outgroup members.
2. Yamagishi and Mifune (2008) did not name the hypothesis, but Mifune (2011) did in Japanese.
3. We used the means and standard deviations of cooperation from Imada et al.’s (2023) pilot study.
4. When planning Studies 1 and 2, we were focused only on testing the reputation management hypothesis. As such, we did not include any hypotheses regarding expected cooperation in the preregistration.

5. We also tested separate mediation models, each with a single mediator (either reputational concern or expected cooperation), and found that the overall pattern of effects did not differ from that reported for the parallel mediation model.
6. We also computed this correlation using the data from our Study 2. The correlation between reputational concern and expected cooperation in the ingroup–common, outgroup–common, ingroup–unilateral, and outgroup–unilateral conditions were .01 ( $p = 0.914$ ), .007 ( $p = 0.387$ ), .20 ( $p = 0.018$ ), and .09 ( $p = 0.286$ ), respectively.

### References

- Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556–1581. <https://doi.org/10.1037/a0037737>
- Columbus, S., Thielmann, I., Zettler, I., & Böhm, R. (2023). Parochial reciprocity. *Evolution and Human Behavior*, *44*(2), 131–139. <https://doi.org/10.1016/j.evolhumbehav.2023.02.001>
- Everett, J. A. C., Faber, N. S., & Crockett, M. (2015a). Preferences and beliefs in ingroup favoritism. *Frontiers in Behavioral Neuroscience*, *9*, 1–21. <https://doi.org/10.3389/fnbeh.2015.00015>
- Everett, J. A. C., Faber, N. S., & Crockett, M. J. (2015b). The influence of social preferences and reputational concerns on intergroup prosocial behaviour in gains and losses contexts. *Royal Society Open Science*, *2*(12). <https://doi.org/10.1098/rsos.150546>
- Ge, X. (2023). Experimentally manipulating mediating processes: Why and how to examine mediation using statistical moderation analyses. *Journal of Experimental Social Psychology*, *109*, Article 104507. <https://doi.org/10.1016/j.jesp.2023.104507>
- Hackel, L. M., Zaki, J., & van Bavel, J. J. (2017). Social identity shapes social valuation: Evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience*, *12*(8), 1219–1228. <https://doi.org/10.1093/scan/nsx045>
- Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(3859), 1243–1248. <https://doi.org/10.1126/science.162.3859.1243>
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, *78*(1), 81–91. <http://doi.org/10.1037//0022-3514.78.1.81>

- Imada, H. (2023). The relative effectiveness of positive and negative gossip in promoting prosocial giving: The examination of the valence of gossip content and reputational consequences. *Japanese Psychological Research*. Advance online publication. <https://doi.org/10.1111/jpr.12473>
- Imada, H., Codd, D., & Liu, D. (2021). Intergroup discrimination in cooperation among moral and non-moral groups. *Letters on Evolutionary Behavioral Science*, 12(1), 28–33. <https://doi.org/10.5178/LEBS.2021.86>
- Imada, H., Hopthrow, T., & Abrams, D. (2021). The role of positive and negative gossip in promoting prosocial behavior. *Evolutionary Behavioral Sciences*, 15(3), 285–291. <https://doi.org/10.1037/EBS0000218>
- Imada, H., Romano, A., & Mifune, N. (2023). Dynamic indirect reciprocity; when is indirect reciprocity bounded by group membership? *Evolution and Human Behavior*, 44(1), 60–66. <https://doi.org/10.1016/j.evolhumbehav.2022.12.001>
- Jin, N., Yamagishi, T., & Kiyonari, T. (1996). Bilateral dependency and the minimal group paradigm. *Japanese Journal of Psychology*, 67(2), 77–85. <https://doi.org/10.4992/JJPSY.67.77>
- Kajiwara, T., Myowa, M., & Mifune, N. (2022). Trait negative reputational concerns among in-group members and in-group favoritism in minimal groups. *Letters on Evolutionary Behavioral Science*, 13(1), 6–9. <https://doi.org/10.5178/LEBS.2022.91>
- Karp, D., Jin, N., Yamagishi, T., & Shinotsuka, H. (1993). Raising the minimum in the minimal group paradigm. *The Japanese Journal of Experimental Social Psychology*, 32(3), 231–240. <https://doi.org/10.2130/jjesp.32.231>
- Koopmans, R., & Rebers, S. (2009). Collective action in culturally similar and dissimilar groups: An experiment on parochialism, conditional cooperation, and their linkages. *Evolution and Human Behavior*, 30(3), 201–211. <https://doi.org/10.1016/j.evolhumbehav.2009.01.003>
- Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1). <https://doi.org/10.1177/2515245920951503>
- Lazić, A., Purić, D., & Krstić, K. (2021). Does parochial cooperation exist in childhood and adolescence? A meta-analysis. *International Journal of Psychology*, 56(6), 917–933. <https://doi.org/10.1002/ijop.12791>
- Leonardelli, G. J., & Brewer, M. B. (2001). Minority and majority discrimination: When and why. *Journal of Experimental Social Psychology*, 37(6), 468–485. <https://doi.org/10.1006/jesp.2001.1475>
- Mifune, N. (2011). *In-group altruism as a reputation management strategy* [PhD thesis, Hokkaido University]. <https://ci.nii.ac.jp/naid/500000551551?l=en>
- Mifune, N., Hashimoto, H., & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. *Evolution and Human Behavior*, 31(2), 109–117. <https://doi.org/10.1016/j.evolhumbehav.2009.09.004>
- Mifune, N., & Yamagishi, T. (2015). A test of the correlation between ingroup favoritism and fear of negative evaluation. *Japanese Journal of Social Psychology*, 31(2), 128–134. [https://doi.org/10.14966/jssp.31.2\\_128](https://doi.org/10.14966/jssp.31.2_128)
- Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the “tragedy of the commons.” *Nature*, 415(6870), 424–426. <https://doi.org/10.1038/415424a>
- Muthén, L. K., and Muthén, B. O. (1998–2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291–1298. <https://doi.org/10.1038/nature04131>
- Otten, S. (2016). The minimal group paradigm and its maximal impact in research on social categorization. *Current Opinion in Psychology*, 11, 85–89. <https://doi.org/10.1016/j.copsyc.2016.06.010>
- Rabbie, J. M., & Horwitz, M. (1988). Categories versus groups as explanatory concepts in intergroup relations. *European Journal of Social Psychology*, 18(2), 117–123. <https://doi.org/10.1002/ejsp.2420180204>
- Rabbie, J. M., Schot, J. C., & Visser, L. (1989). Social identity theory: A conceptual and empirical critique from the perspective of a behavioural interaction model. *European Journal of Social Psychology*, 19(3), 171–202. <https://doi.org/10.1002/ejsp.2420190302>
- Romano, A., Balliet, D., & Wu, J. (2017). Unbounded indirect reciprocity: Is reputation-based cooperation bounded by group membership? *Journal of Experimental Social Psychology*, 71, 59–67. <https://doi.org/10.1016/j.jesp.2017.02.008>
- Romano, A., Sutter, M., Liu, J. H., Yamagishi, T., & Balliet, D. (2021). National parochialism is ubiquitous across 42 nations around the world. *Nature Communications*, 12(1), 1–8. <https://doi.org/10.1038/s41467-021-24787-1>



- Simpson, B. (2006). Social identity and cooperation in social dilemmas. *Rationality and Society*, 18(4), 443–470. <https://doi.org/10.1177/1043463106066381>
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178. <https://doi.org/10.1002/ejsp.2420010202>
- Tajfel, H., & Turner, J. (1979). An integrative theory of inter-group conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–37). Brooks/Cole.
- Thomson, R., Yuki, M., Talhelm, T., Schug, J., Kito, M., Ayanian, A. H., Becker, J. C., Becker, M., Chiu, C. Y., Choi, H. S., Ferreira, C. M., Fülöp, M., Gul, P., Houghton-Illera, A. M., Joasoo, M., Jong, J., Kavanagh, C. M., Khutkyy, D., Manzi, C., . . . Visserman, M. L. (2018). Relational mobility predicts social behaviors in 39 countries and is tied to historical farming and threat. *Proceedings of the National Academy of Sciences of the USA*, 115(29), 7521–7526. <https://doi.org/10.1073/pnas.1713191115>
- Waller, I., & Anderson, A. (2021). Quantifying social organization and political polarization in online platforms. *Nature*, 600, 264–268. <https://doi.org/10.1038/s41586-021-04167-x>
- Weisel, O., & Böhm, R. (2015). “Ingroup love” and “outgroup hate” in intergroup conflict between natural groups. *Journal of Experimental Social Psychology*, 60, 110–120. <https://doi.org/10.1016/j.jesp.2015.04.008>
- Wu, J., Balliet, D., & van Lange, P. A. M. (2015). When does gossip promote generosity? Indirect reciprocity under the shadow of the future. *Social Psychological and Personality Science*, 6(8), 923–930. <https://doi.org/10.1177/1948550615595272>
- Wu, J., Balliet, D., & van Lange, P. A. M. (2016). Reputation, gossip, and human cooperation. *Social and Personality Psychology Compass*, 10(6), 350–364. <https://doi.org/10.1111/spc3.12255>
- Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity: Ingroup boasting and ingroup favouritism. *Advances in Group Processes*, 16, 161–197.
- Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, 63(2), 116–132. <https://doi.org/10.2307/2695887>
- Yamagishi, T., & Mifune, N. (2008). Does shared group membership promote altruism? Fear, greed, and reputation. *Rationality and Society*, 20(1), 5–30. <https://doi.org/10.1177/1043463107085442>
- Yamagishi, T., & Mifune, N. (2016). Parochial altruism: Does it explain modern human group psychology? *Current Opinion in Psychology*, 7, 39–43. <https://doi.org/10.1016/j.copsyc.2015.07.015>
- Yamagishi, T., Mifune, N., Liu, J. H., & Pauling, J. (2008). Exchanges of group-based favours: Ingroup bias in the prisoner’s dilemma game with minimal groups in Japan and New Zealand. *Asian Journal of Social Psychology*, 11(3), 196–207. <https://doi.org/10.1111/j.1467-839X.2008.00258.x>