Evolution of group-wise cooperation: Is direct reciprocity insufficient?

Shun Kurokawa[a,b,c,*], Yasuo Ihara[a]

[a] Department of Biological Sciences, the University of Tokyo, Tokyo 113-0033, Japan
[b] Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan
[c] Institute of Zoology, Chinese Academy of Sciences, Datun Road, Chaoyang, Beijing 100101, PR China

ABSTRACT

Group-wise cooperation, or cooperation among three or more individuals, is an integral part of human societies. It is likely that group-wise cooperation also played a crucial role in the survival of early hominins, who were confronted with novel environmental challenges, long before the emergence of *Homo sapiens*. However, previous theoretical and empirical studies, focusing mainly on modern humans, have tended to suggest that evolution of cooperation in sizable groups cannot be explained by simple direct reciprocity and requires some additional mechanisms (reputation, punishment, etc.), which are cognitively too demanding for early hominins. As a partial resolution of the paradox, our recent analysis of a stochastic evolutionary model, which considers the effect of random drift, has revealed that evolution of group-wise cooperation is more likely to occur in larger groups when an individual's share of the benefit produced by one cooperator does not decrease with increasing group size (i.e., goods are non-rivalrous). In this paper, we further extend our previous analysis to explore possible consequences of introducing rare mistakes in behavior or imperfect information about behavior of others on the model outcome. Analyses of the extended models show that evolution of group-wise cooperation can be facilitated by large group size even when individuals intending to cooperate sometimes fail to do so or when all the information about the past behavior of group members is not available. We argue, therefore, that evolution of cooperation in sizable groups does not necessarily require other mechanisms than direct reciprocity if the goods to be produced via group-wise cooperation are non-rivalrous.
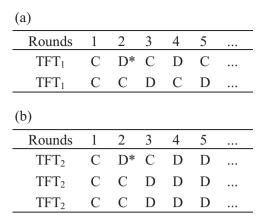
1. Introduction

Group-wise cooperation, or cooperation involving more than two individuals, is likely to have played an indispensable role in the survival and prosperity of human societies (Boyd and Richerson, 1988; Joshi, 1987). A vast number of models have been proposed to understand the evolution of cooperation in terms of natural selection (Axelrod and Hamilton, 1981; Hamilton, 1964; Nowak, 2006; Trivers, 1971). Although the evolutionary origin of human group-wise cooperation is by no means clear, one possibility is that it emerged in an earliest stage of human evolution, after the divergence from the chimpanzee lineage. The hypothesis is worth considering because group-wise cooperation might have been a crucial factor that enabled early hominins to survive in the open savanna environment, which they newly occupied. Hominins in the first few million years since the divergence can be properly called "bipedal apes," with their brains being comparable in size with those of extant chimpanzees. It follows, therefore, that if group-wise cooperation indeed played a role in the survival of early hominins, it must have been realized by demanding only moderate cognitive capacities, not as those of *Homo sapiens*.

Reciprocity is a key concept to explain the evolution of cooperation between a pair of unrelated individuals (Axelrod, 1984; Axelrod and Hamilton, 1981; Trivers, 1971). The logic is, in a nutshell, as follows: given the prospect that an ongoing interaction with a partner will continue for sufficiently long, it pays for an individual to engage in costly cooperative behavior to benefit the partner because that encourages the partner's future cooperation. As for the evolution of group-wise cooperation, however, many authors have claimed that direct reciprocity is insufficient and some additional mechanisms are needed (Boyd and Richerson, 1988; Boyd et al., 2003; Fehr, 2004; Fehr and Fischbacher, 2003; Hagen and Hammerstein, 2006; Hauert et al., 2002; Henrich, 2004; Joshi, 1987; Kurzban and Houser, 2005; Hilbe et al., 2014). There are three rationales for that view.

First, evolution of cooperation is possible only if a defection is somehow retaliated. In the context of direct reciprocity, an individual retaliates against a defector by withdrawing future cooperation. In dyadic interactions, this kind of retaliation is effective, because it is focused on the very individual who is uncooperative. In group-wise

(a)

Rounds	1	2	3	4	5	...
TFT_1	C	D*	C	D	C	...
TFT_1	C	C	D	C	D	...

(b)

Rounds	1	2	3	4	5	...
TFT_2	C	D*	C	D	D	...
TFT_2	C	C	D	D	D	...
TFT_2	C	C	D	D	D	...

Fig. 1. Sample behavioral sequences of TFT_as in a group of size n for (a) $n=2$, $a = 1$ and (b) $n = 3$, $a = 2$. In both cases, a TFT_a fails to cooperate by mistake, as indicated by the asterisk, in the second round.

interactions, on the contrary, withdrawal of cooperation harms all individuals in the group, some of whom may be cooperative, and thus less effective as a way to punish a defector (Hauert et al., 2002). As detailed below, a mathematical analysis by Boyd and Richerson (1988) has provided support for this argument.

Second, rare mistakes in behavior, such that an individual who intends to cooperate sometimes fails to do so, may affect the evolution of cooperation (May, 1987; Kurokawa, 2016c). There are reasons to suspect that mistakes in behavior may have a greater impact on the evolution of cooperation in group-wise interactions than in the dyadic counterpart. Obviously, the expected number of mistakes increases linearly with the group size. In addition, more importantly, group-wise cooperation is vulnerable to a "chain reaction" triggered by a mistake, from which dyadic cooperation is partially immune. Let us illustrate the point using a simple example. Consider n-player repeated games and a strategy called TFT_a, which cooperates in the first round and then cooperates in each subsequent round if a or more of the $n - 1$opponents cooperated in the previous round (Taylor, 1976). Imagine, on the one hand, a game between a pair of individuals adopting TFT_1 ($n = 2$; Fig. 1a) and a game among a trio of TFT_2 individuals ($n = 3$; Fig. 1b), on the other. The chain reaction in dyadic interactions is characterized by a series of unilateral cooperation with alternation of roles played by the two individuals, whereas in interactions involving more than two individuals, the chain reaction results in a full collapse of cooperation (Nowak and Sigmund, 1992). More formally, a simulation study by Bowles and Gintis (2011) has suggested that direct reciprocity does not support cooperation in groups of size larger than two unless mistakes are virtually absent.

Third, it has been pointed out that information structure among individuals may have a significant effect on the evolution of cooperation (Bowles and Gintis, 2011; Kurokawa, 2016a, 2016b, 2016d). In direct reciprocity, in particular, information about how an individual has behaved previously is of prime importance. This is because for an act of cooperation to serve to signal one's willingness to cooperate and thereby elicit cooperation by others, it has to be known by them. Hence, evolution of cooperation by direct reciprocity is generally discouraged by imperfect information. More to the point, the number of acts that are not recognized accurately by group members is expected to increase with the group size, which is another basis for the claim that direct reciprocity is insufficient to explain the evolution of cooperation in sizable groups. Not surprisingly, Bowles and Gintis's (2011) simulation has suggested that evolution of group-wise cooperation by direct reciprocity is unlikely without information being perfect.

In addition to these theoretical accounts, there is also an empirical study that supports the notion that cooperation is less likely in larger groups. In one experiment, school children played n-player repeated prisoners' dilemma games, where the children were provided with candy bars and asked whether to donate them to the group (Alencar et al., 2008). The donated candy bars were tripled and then equally divided among all the children. They found that group size affects the level of cooperation, with children in large groups cooperating significantly less than those in small groups.

If it is true that direct reciprocity cannot by itself account for the evolution of human group-wise cooperation, then some additional explanation needs to be sought out. Accordingly, a number of candidate mechanisms have been proposed in the last several decades; for example, genetic or cultural group selection (Boyd et al., 2003; Henrich, 2004; Wilson and Sober, 1994), punishment (Boyd and Richerson, 1992), reward (Rand et al., 2009), indirect reciprocity (Panchanathan and Boyd, 2004; Suzuki and Akiyama, 2008), costly signaling (Gintis et al., 2001), opting out of a group (Hauert et al., 2007), policing (Frank, 1995), and strong reciprocity (Gintis, 2000). A problem of these additional mechanisms is that they seem to demand considerably higher cognitive capacities than direct reciprocity. We believe that these are plausible models of what might have occurred after the emergence of *Homo sapiens*, but unlikely to account for group-wise cooperation in early hominins, which we consider here.

However, we argue, these are only half of the story. That is because most of the above arguments against direct reciprocity as an explanation for the evolution of group-wise cooperation assume implicitly or explicitly that the benefit produced by cooperators is non-excludable and rivalrous (see Dionisio and Gordo, 2006). Excludability concerns whether or not any individuals can be excluded from use of the benefit and rivalry is about whether use of the benefit by one individual reduces its availability to others. In the context of theoretical and empirical games, a benefit is non-excludable if (but not only if) it is divided equally among all players. A benefit is rivalrous if each player's share of the benefit produced by a single cooperator decreases with increasing number of players. A good that is non-excludable and rivalrous is called a common good (Taylor, 1976).

Our point is that the benefit produced by cooperators is not necessarily rivalrous and if we consider a non-rivalrous benefit the above argument for the insufficiency of direct reciprocity loses its ground. An example of cooperative behavior that produces non-rivalrous benefit is guarding and policing of group territory. Since these behaviors tend to increase the likelihood that the group survives between-group conflicts, all group members enjoy the benefit (i.e., non-excludable) and the amount of benefit obtained by each member does not decrease with increasing number of members (i.e., non-rivalrous) as long as the size, and thus the defensibility of the territory is constant. Another example may be knowledge of innovations shared by group members (Hess and Ostrom, 2006). In humans and perhaps in some non-human animals as well, innovations that enhance individual's fitness can be socially transmitted among group members to be used freely. Thus, an act of costly innovation can be regarded as a form of cooperation that produces non-excludable and non-rivalrous benefit. A good that is non-excludable and non-rivalrous is referred to as a public good (Taylor, 1976).

The effect of group size on the evolution of group-wise cooperation has been formally investigated by Boyd and Richerson (1988). Importantly, they considered both when the benefit produced by cooperators is rivalrous and when it is non-rivalrous. They analyzed deterministic evolutionary dynamics of reciprocators and defectors in the n-player prisoners' dilemma game. Two equilibrium states always exist, one of which is the cooperative equilibrium, where every individual in the population is a reciprocator, and the other is the non-cooperative equilibrium, where no reciprocator exists. The cooperative equilibrium can be stable against small perturbation given that games are repeated sufficiently many times before a group is dismissed. Even when the cooperative equilibrium is locally stable, however, the non-cooperative equilibrium is always locally stable, too, so that which of the two equilibria are eventually reached depends on the initial state of the population. Boyd and Richerson (1988) showed that the basin of