強い互酬性仮説の検証¹

——協力行動と罰行動の関係——

李 楊² 北海道大学 山岸 俊男 東京大学

A test of the strong reciprocity model: Relationship between cooperation and punishment

Yang Li (*Hokkaido University*) and **Toshio Yamagishi** (*University of Tokyo*)

The strong reciprocity model of human cooperation (SRM) argues that strong reciprocators, who cooperate with others and punish non-cooperators, sustain within-group cooperation. However, the assumption that altruism and punishment are products of the same psychological mechanism of strong reciprocity has not been fully verified. Second-party punishment, for example as measured through rejection of unfair offers in the ultimatum game, has been demonstrated to have no relationship with cooperation in the prisoner's dilemma and other games. In this study, we tested the assumption of the SRM by comparing the participants' levels of cooperation in the prisoner's dilemma game and their inclination for punishment in a third-party punishment game. Non-student recruited from the general population ($N = 182$) participated in the study. The results show a weak but positive correlation between cooperation and third-party punishment, which is consistent with the SRM model.

Key words: strong reciprocity model, altruism, punishment, prisoner's dilemma game, third-party punishment game.

The Japanese Journal of Psychology
2014, Vol. 85, No. 1, pp. 100–105

ヒトの個体が他の動物種には見られない強い向社会性を備えていることは広く認められている (Bowles & Gintis, 2011) が、なぜヒトがそうした強い向社会性を備えるに至ったかは進化の大きな謎のひとつである。現在、こうしたヒトの向社会性を説明するための理論的研究は理論生物学と社会科学の双方で進められており、大きく分けて二つの理論的可能性が追究されている。一つは向社会行動が他者からの向社会行動を引き出し、その結果引き出される利益が向社会行動を取るためのコストを上回るという、間接互酬性モデルに代表される広い意味での利益還元理論である (Nowak &

Sigmund, 2005)。もう一つは、向社会行動はそうした行動を取る個体に利益をもたらさないが、向社会行動を取る個体が多くいる集団に属する個体は、そうした個体が少ない集団に属する個体よりも平均して大きな利益を獲得するため向社会性が進化可能となるとする、文化的群淘汰モデルを含む広い意味での群淘汰理論 (Boyd & Richerson, 1985; Gintis, 2000) である。本論文で検討する強い互酬性モデル (strong reciprocity model: 以下 SRM とする) (Fehr & Fischbacher, 2003; Gintis, 2000) は後者の理論の一部であり、ヒトは集団全体の利益を向上させる正の互酬性 (他人が協力している限り自分も協力する) と同時に、そうした正の互酬性による行動を安定させるための負の互酬性 (協力しない人間を罰する傾向) も進化を通して獲得したとする理論である (Bowles, Choi, & Astrid, 2003; Bowles & Gintis, 2004; Boyd, Gintis, Bowles, Richerson, 2003; Choi & Bowles, 2007; Gintis, 2000; Henrich & Boyd, 2001)。集団利益に反する規範逸脱者を罰することで集団内での利益の平準化をもたらす傾向をヒトが備えていれば、集団間の淘汰圧は相対的に上昇し、集団内での淘汰圧が弱くなるため、群淘汰が進化に対して及

Correspondence concerning this article should be sent to: Yang Li, Hokkaido University, Graduate School of Letters, Faculty of Letters, Hokkaido University, Kita-ku, Sapporo 060-0810, Japan (e-mail: lee@lynx.hokudai.ac.jp)

¹ 本研究は JSPS 科研費 19046005 の助成を受けた。

² 本研究の実施にあたりご協力いただいた清成 透子氏、谷田 林士氏、品田 瑞穂氏、三船 恒裕氏、高岸 治人氏、堀田 結孝氏、橋本 博文氏、高橋 知里氏、犬飼 佳吾氏、波多野 礼佳氏、三浦 亜利紗氏、Dora Simunovic 氏、森 康浩氏、山崎 圭子氏に深く感謝いたします。

ばす力が強くなる。そのため、強いSRMは群淘汰理論によるヒトの向社会性の説明において一定の役割を果たしている。

SRMの提唱者たちは、モデルの検証のために一連の実験研究を用いている。そのために主として用いられてきたのは、罰つき公共財ゲーム実験、最後通告ゲーム実験、および第三者罰実験である。Yamagishi (1986) が開始した罰つき公共財ゲーム実験では、参加者は公共財に対し協力するかどうかを決めると同時に、協力しないメンバーに対して罰を与えるかどうかを決定する。SRMによれば、罰つき公共財ゲームで非協力者に対して罰を与えるのは協力者のはずである (Fehr & Schmidt, 1999)。この予測は、Kiyonari & Barclay (2008) の実験で、公共財ゲームでの協力者は非協力者よりも非協力者を罰する傾向が強いという結果により支持されている。またYamagishi & Takahashi (1994) は、コンピュータシミュレーションを用いて、公共財ゲームの参加者が強い互酬性に従って行動する——協力者が非協力者を罰する傾向を持つ——場合には、相互協力が成立し公共財問題の解決が可能となることを示している。ただし協力行動と罰行動の間の一貫性が、強い互酬性に基づいて生み出されるという証拠はない。強い互酬性とは独立の、例えばフラストレーション・アグレッション (Dollard, Doob, Miller, Mowrer, & Sears, 1939 宇津木訳 1959) ——相互協力を目指して協力したのに相互協力が達成できないという目標の達成の阻害により生じるフラストレーションが、目標の達成を妨げている非協力者に対する攻撃行動を生み出す——によるものである可能性も残されている。従って、SRMで想定される心理傾向を人間が備えていることを検証するためには、同一人物が異なるゲーム——例えば公共財ゲームなどの協力行動を測定するためのゲームと、非協力者を罰する第三者罰ゲーム——に異なる機会に参加し、それぞれのゲームで独立に決定した行動間に一貫性が見られるかどうかを調べる必要がある。

異なる実験ゲーム間で協力行動と罰行動が一貫するかどうかを検討した研究としては、Yamagishi, Horita, Mifune, Hashimoto, Li, Shinada, Miura, Inukai, Takagishi, & Simunovic (2012) による、囚人のジレンマゲームと最後通告ゲームとの関連を調べた研究がある。最後通告ゲームにおける不公平提案への拒否行動は、これまで強い互酬性仮説を支持する強力な証拠、つまり公平性に反する行動を取る人間を罰する傾向を人間が有していることの証拠として用いられてきた (Fehr & Fischbacher, 2003; Fehr, Fischbacher, & Gächter, 2002; Gintis, Boyd, Bowles, & Fehr, 2003)。ただし、これまでの研究では、不公平提案を拒否する人間がある比率で存在するという事実のみにもとづいて、人々は公平性の規範に違反する他者をコストをかけて罰する傾向を

持っていることの証拠とされてきた。これに対してYamagishi et al. (2012) は、人々が実際にSRMと一貫する行動傾向を持っているとすれば、社会的ジレンマや囚人のジレンマで協力する傾向の強い人は、最後通告ゲームにおいて不公平提案を拒否する傾向が強いはずだという仮説を検証するための研究を行った。この研究では、それぞれ最低数ヶ月の間隔を置いて同一参加者に最後通告ゲーム、囚人のジレンマゲーム、独裁者ゲーム、信頼ゲームに参加してもらい、最後通告ゲームにおいて不公平提案を拒否する傾向が、囚人のジレンマゲームにおける協力傾向や、信頼ゲームにおいて信頼にこたえる傾向、あるいは独裁者ゲームにおける利他行動と正の相関を持つかどうかを検討した。その結果、それらの間にまったく相関が存在しない、あるいは囚人のジレンマゲームの場合には弱い有意な負の相関が存在することが明らかにされた。Yamagishi et al. (2012) はこの結果に基づき、少なくとも最後通告ゲームにおける不公平提案の拒否は、強い互酬性に従う傾向を人々が有していることの証拠にはなりえないという結論を下している。またこのYamagishi et al. (2012) に対するコメントの中でEgloff, Richter, & Schmukle (2013) は、ドイツにおける大規模調査 (Socio-Economic Panel 調査 version 28) の結果に基づき、利他行動を取る人間に対してポジティブな反応を取る傾向と、公平性の違反に対して罰を与える傾向との間に関連性が見られないことを明らかにしている。

このように、社会関係において協力的な人間は規範逸脱者に対して罰行動を取る傾向が強いというSRM仮説に対しては、これまで一貫した実験結果が示されていない。本研究の目的は、Yamagishi et al. (2012) の研究では報告されていない第三者罰 (TPP) ゲームを用いて、囚人のジレンマでの協力傾向と、独裁者ゲームで利益を独り占めした人に対して罰を与える傾向とが一貫するかどうかを検討することにある。自分とは関係のない他者に対して不公正な行動を取る第三者に対する罰は、最後通告ゲームにおける不公平提案の拒否と同様、Fehr & Fischbacher (2004) などのこれまでの研究で、強い互酬性傾向の証拠として用いられてきた。また、協力傾向の指標として囚人のジレンマでの協力度に加え、社会的価値志向性 (Liebrand, 1984; Liebrand & McClintock, 1988) の測定を行い、第三者罰傾向の強さと社会的価値志向性における向社会性との間に関連が見られるのかも検討した。

方 法

本研究は、Yamagishi et al. (2012) の研究で用いられた“ゲームと文化”データセットに、新たに第三者罰実験を追加するかたちで実施した。参加者は、2007年12月に札幌市北区を中心に配布された新聞の折り込み広告を用いて募集した。募集広告には、(a) 北海

道大学で実施する心理学実験の参加者を募集すること、(b) 学生は応募できないこと、(c) 実験参加者には報酬が支払われること、(d) 実験は何度も繰り返されるため、なるべく継続的な参加が望まれることなどが記されていた。折り込み広告に対する応募者のうちから男女比が偏らないようにするかたちで 108 名 (男性 51 名, 女性 57 名)³が参加者として選ばれた (以下, 第 1 サンプルとする)。2008 年 2 月から 2010 年 8 月にかけて、これらの参加者に対して計 6 回にわたり調査・実験を行い、各回に参加者に対して各種の実験への参加と質問紙への回答を求めた。また 2011 年 2 月には上述とはほぼ同様の方法で新たに募集した 105 名 (男性 52 名, 女性 53 名) の参加者 (以下, 第 2 サンプルとする) を用いた研究が実施され、同年 11 月には上述の二つのサンプルを合わせた 184 名 (男性 90 名, 女性 94 名) を用いた研究が実施されている。本研究では、2011 年 11 月に実施した第三者罰ゲーム実験の結果を用いて分析を行う。分析に用いた 182 名の参加者の 2011 年 2 月における平均年齢は 46.5 歳、最も若い参加者の年齢は 23 歳、最高齢の参加者は 70 歳であった。

第三者罰ゲーム

本研究での第三者罰ゲームはシナリオ実験として行われ、他の二者の間で不公平分配が行われた際に、不公平分配を行った人の利得を減らすために参加者自身が支払っても良いと思うコストの程度を、第三者罰傾向の指標とした。参加者には、他の 2 人が行っている実験の様子を自分が観察している状況を想定してもらった。その際、参加者は 2,000 円の報酬を与えられているものとした。観察している実験は別の 2 人の参加者の間で行われた独裁者ゲームであり、そのうちの 1 人 (“分配者”) が実験者から 3,000 円を元手として与えられ、その 3,000 円をもう 1 人の参加者 (“受け手”) との間で自由に分配するというものであった。シナリオではさらに、独裁者ゲームでの分配者は、実験者から与えられた 3,000 円をすべて自分のものにし、受け手に 1 円も分配しなかった場面が記されていた。その上で、参加者には、自分の報酬の中から支払う金額の 2 倍の金額を分配者の利得から減らすことができるとしたら、参加者は 3,000 円を独り占めした (シナリオ上の) 分配者の利得を減らすためにいくら使いたいかを回答してもらった。回答は、罰に使用する金額を 1,500 円, 1,000 円, 500 円, 300 円, 100 円, 0 円の六つの選択肢から選ぶ形で行われた。

囚人のジレンマゲーム

参加者は囚人のジレンマゲームを 2 回にわたり経験した。2 回目のゲーム (PD2) は TPP 実験と同じ 2011 年 11 月に全サンプルが参加して実施されたが、1 回目のゲーム (PD1) は第 1 サンプルに対しては 2009 年 2 月に、第 2 サンプルに対しては 2011 年 2 月に、同じ手順を用いて実施された。

PD1 は、一度限りの同時決定囚人のジレンマゲームとして実施された。参加者は同一セッションに参加した他の 5 人から 9 人の参加者のいずれかを相手に “お金のやり取り” を行った。この “やり取り” では、まず各参加者に 2,000 円の元手が与えられた。そして参加者は、そのうちのいくらを相手に渡すかを 100 円単位で決定するように求められた。また同時に相手も同じ決定を行っていることが教示された。お互い相手に渡す金額は 2 倍にされた後相手の利得に加算され、渡さない金額はそのまま自分の利得に加算された。すなわち、この囚人のジレンマゲームにおける参加者の報酬は、相手に渡さないで自分の手元に残しておいた金額と、相手から渡された金額の 2 倍の合計であった。ゲームの決定内容をより直感的に理解してもらうため、各参加者には決定用紙とともに 20 枚の “100 円券” が渡され、相手に渡す金額に対応する枚数の “100 円券” を封筒に “決定用紙” とともに入れるように教示された。それぞれの参加者はこれらの作業を個人ごとの個室で行った。

PD2 の参加者は、PD1 と同様の一度限りの同時決定囚人のジレンマゲームを、毎回異なる相手との間で計 3 回行った。ただし参加者には、毎回の相手が誰であるか、またゲームを何回行うかは知らされていなかった。ゲームは 100 円券と封筒を用いてではなく、個室におかれたコンピューターを用いて行われた。それぞれの回ごとに参加者は 500 円の元手を与えられ、ランダムに組み合わせられた他の 1 人の参加者に対してその元手のうちいくらを相手に渡すかを 50 円単位で決定した。決定に際しては、相手も同時に同じ決定を行うことが教示されていた。それぞれの回で参加者は、自分が手元に残しておいた金額と、ランダムにマッチされたもう 1 人の参加者が提供した金額の 2 倍の合計を獲得したが、相手の行動が 2 回目以降の行動に対して与える影響を除くため、相手の行動および獲得金額のフィードバックは各試行終了時には行わず、すべてのゲーム終了後に獲得金額の合計を支払うかたちで行われた。

社会的価値志向性

囚人のジレンマゲームにおける協力的行動のほかに、参加者の向社会的性の指標として参加者の社会的価値志向性 (Liebrand, 1984; Leibrand & McClintock, 1988) を

³ 分析では回答に一貫性のない参加者 1 名と実験内容への理解が著しく低い参加者 1 名を除いた 106 名 (男性 51 名, 女性 55 名) が用いられている。

測定した。測定はリング法 (Liebrand & McClintock, 1988) を用いて 2011 年 11 月に全サンプルを対象に実施された。具体的には、参加者は見ず知らずのもう 1 人の人と自分がそれぞれお金をもらわないし失う状況を想定したうえで、自分と相手の利得が示されている二つの選択肢 (例：選択肢 A では自分は 390 円を手に入れ、相手は 1,450 円を失う、選択肢 B では自分は 0 を手に入れ、相手は 1,500 円を失う) のどちらかを選択する質問を 24 問回答した。これらの質問に対する回答を集計し、自己利益を重視する程度と相手の利益を重視する程度を比較することで、相手の利益を無視して自分の利益のみを重視する独立主義者 ($n = 88$)、相手の利益を自分の利益と同程度に重視する協

力主義者 ($n = 51$)、相手の利益を自分の利益よりも優先する利他主義者 ($n = 7$)、相手の利益を下げることを重視する競争主義者 ($n = 3$) に分類したうえで、近年の分類法に従い、独立主義者と競争主義者を向自己的価値志向性の持ち主 ($n = 91$)、協力主義者と利他主義者を向社会的価値志向性の持ち主 ($n = 58$) として分類した。残りの 33 名の参加者はいずれのカテゴリにも分類不能とされたため、社会的価値志向性に関する分析からは外した。

結 果

ゲーム行動

第三者罰ゲームにおける罰金額の分布を Figure 1 に示す。罰額の平均は 318.68 円 ($SD = 460.67$) であった。囚人のジレンマゲームでの協力行動 (元手に対する供出額の比率) は、平均して PD1 では元手 2,000 円のうち 44% ($SD = 32.88$) を供出し、PD2 では元手 500 円の 38% ($SD = 31.82$) を供出した。

第三者罰との相関

第三者罰ゲームで罰のために使用した金額と PD1 と PD2 における協力度は、それぞれ r_s (スピアマンの順位相関係数) = 0.19 ($p < .05$) と 0.23 ($p < .01$) であった⁴。PD1 と第三者罰の金額の散布図を Figure 2 に示す。PD1 と第三者罰との間には最低ほぼ半月の間隔があったが、PD2 との相関とそれほど変わらない関係が見られた。また向社会的性と PD1, PD2, 罰額との間にはそれぞれ $r_s = 0.35$ ($p < .0001$)、0.39 ($p < .0001$)、および 0.26 ($p < .001$) の相関が見られた。これらの結果は、協力傾向と罰傾向との間に一貫性が有意に見られることを示している。また、罰額と向社会的価値志向性との間の点双列相関係数は 0.30 ($p < .001$)⁵である。この結果は、罰行動が (協力行動を維持するための) 2 次的協力行動であることを示している。

Yamagishi et al. (2012) では、第 1 サンプルのみを用いて実施された最後通告ゲームでの不公平提案に対する拒否行動が、PD2 と弱い有意な負の相関にあることが報告されている。本研究では追加分析として、第 1 サンプルのみを用いて最後通告ゲームにおける不公平提案と第三者罰ゲームの関連の分析を行った。その結果、不公平提案に対する拒否行動と第三者ゲームにおける罰行動の間には有意な相関は見られなかった ($r_s = -0.07, ns$)。

Figure 1. Proportion of third-party punishment.

Figure 2. Correlation between cooperation in PDG1 and third-party punishment.

Note. Area of the circles stands for number of subjects.

⁴ ピアソン相関係数は 0.23 (PD1, $p < .005$) および 0.19 (PD2, $p < .01$) であった。

⁵ 社会的価値志向性の向社会的と向自己的との間の第三者罰の平均はそれぞれ 432.8 円 ($SD = 556.4$) と 154.9 円 ($SD = 334.1$) であり、有意な差が見られる ($t(83.4) = 3.43, p < .001$)。

考 察

本研究では第三者罰の傾向と協力行動との間の関係を検討した結果、第三者罰の傾向が囚人のジレンマゲームでの協力行動と正に相関することが明らかにされた。条件付き協力行動と罰行動とが完全に貫している場合には協力行動の進化が可能であることは、これまで Yamagishi & Takahashi (1994), Lehmann, Rousset, Roze, & Keller (2007) らにより明らかにされているが、強い互酬性仮説の提唱者である Bowles & Gintis (2004) は、罰に基づく協力の進化は協力行動と罰行動の間に強い相関を必要としないという議論を行っており、罰が協力の進化につながる道筋についての理論的な決着はいまだついていない。この理論モデル上の論争は、協力行動と罰行動とが現実にはどの程度強く関連しているかと独立に結論を下すべき問題ではない。それぞれの理論の前提が成り立つかどうかを確認するためには、今後、協力行動と罰行動とがどの程度強く関連しているかについての実証研究の蓄積が求められている。本研究の結果は、これら二つの間に弱い関連性が認められるというものであり、この結果のみでは、いずれの理論的立場を支持するかについては明確な結論を下すことができない。

Yamagishi et al. (2012) は、最後通告ゲームにおける不公正提案に対する拒否行動と囚人のジレンマゲームにおける協力行動との間に弱い負の相関を見出しており、Egloff et al. (2013) はその結果を、強い互酬性そのものを否定する結果として解釈している。今回の実験で第三者罰傾向と協力行動とが有意に正に相関するという結果は、Yamagishi et al. (2012) による最後通告実験結果の Egloff et al. (2013) による解釈には反するものであり、従って Yamagishi et al. (2012) が採用している限定的な解釈、すなわち最後通告ゲームにおける不公正提案の拒否は規範逸脱者に対する罰ではなく、自分が見下されたことに対する反撥であるという解釈にとどめるべきだとする議論の妥当性を示唆している。

第三者罰ゲームにおける罰の動機について、Yamagishi et al. (2012) では第三者罰ゲームにおける罰行動も罰行使者としての自分が軽く扱われないように自己主張する要因が働く可能性があると論じている。もしこの可能性が実際に成立していれば、囚人のジレンマゲームにおける協力行動と第三者罰ゲームにおける罰行動の間に相関は存在しないはずである。しかし通常の第三者罰ゲームでは、分配を行う参加者は罰行使者の存在や自分が罰される可能性があることを念頭に置いた上で分配を行うのに対して、本研究で用いられたシナリオでは、独裁者は罰される可能性があることを念頭に置いた上で分配を行ったと記されていない。よって本研究においては罰行使者が軽く扱われ

ていると感じる可能性は低く、Yamagishi et al. (2012) による解釈の可能性を検証するには適したデザインとはなっていない。第三者罰行動の研究の文脈では今後“軽く扱われないように自己主張する”といった動機が第三者罰行動に与える影響を検討するのに適したデザインを用いる必要があるだろう。

本研究の限界の一つは、罰傾向が場面想定法を用いて測定されていることにある。第三者罰実験では参加者に監視者としての役割を割り振るため、その役割に応じた行動（規範逸脱者を罰する行動）を取りやすい可能性がある。実際の自己利益を含まない場面想定法の場合には更にこの傾向が強められ、その場面に応じた規範に従う行動を取ると答える傾向が強められる可能性がある。すなわち、本実験で得られた結果は、実験者に求められている行動、あるいは規範に従う行動であることが明白である場面想定法を用いたために得られた結果である可能性がある。この点を明らかにするためには、第三者である参加者に対する要求特性の影響を最小限にとどめる形でのデザインを開発し、たまたま目にした他人同士の不公平行動に対して実際にコストをかけて罰を行う傾向があるかどうかを検討すると同時に、その際の罰行動と囚人のジレンマでの協力行動とが関連するかどうかを検討する必要があるだろう。

引用文献

- Bowles, S., Choi, J-K., & Astrid, H. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, **223**, 135–147.
- Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology*, **65**, 17–28.
- Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton, NJ: Princeton University Press.
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of National Academy of Sciences of the United States of America*, **100**, 3531–3535.
- Boyd, R., & Richerson, P. J., (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Choi, J-K., & Bowles, S. (2007). The Coevolution of parochial altruism and war. *Science*, **318**, 636–640.
- Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and aggression*. New Haven: Yale University Press.
- (ドラード, J. 宇津木 保 (訳) (1959). 欲求不満と暴力 誠信書房)
- Egloff, B., Richter, D., & Schmukle, S. C. (2013). Need for conclusive evidence that positive and negative reciprocity and unrelated. *Proceedings of the National Academy of Sciences of the United States of America*, **110** (9) E786; published ahead of print February 4,

- 2013.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, **425**, 785–791.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, **25**, 63–87.
- Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature*, **13**, 1–25.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, **114**, 817–868.
- Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, **206**, 169–179.
- Gintis, H., Boyd, R., Bowles, S., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behaviour*, **24**, 153–172.
- Henrich, J., & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, **208**, 79–89.
- Kiyonari, T., & Barclay, P. (2008). Cooperation in social dilemmas: Free riding may be thwarted by second-order reward rather than by punishment. *Journal of Personality and Social Psychology*, **95**, 826–842.
- Lehmann, L., Rousset, F., Roze, D., & Keller, L. (2007). Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *American Naturalist*, **170**, 21–36.
- Liebrand, W. B. G. (1984). The effect of social motives, communication and group size on behaviour in an N-person multi-stage mixed-motive game. *European Journal of Social Psychology*, **14**, 239–264.
- Liebrand, W. B. G., & McClintock, C. H. (1988). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality*, **2**, 217–230.
- Nowak M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, **437**, 1291–1298.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, **51**, 110–116.
- Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., Miura, A., Inukai, K., Takagishi, H., & Simunovic, D. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 20364–20368.
- Yamagishi, T., & Takahashi, N. (1994). Evolution of norms without metanorms. In U. Schulz, W. Albers, & U. Mueller (Eds.), *Social dilemmas and cooperation*. Berlin: Springer-Verlag. pp. 311–326.

—— 2013. 7. 26 受稿, 2013. 9. 7 受理 ——