

## 科学研究費助成事業 研究成果報告書

平成 27 年 6 月 5 日現在

機関番号：17701

研究種目：研究活動スタート支援

研究期間：2013～2014

課題番号：25885057

研究課題名(和文) リーダーの集中罰による社会的ジレンマの解決

研究課題名(英文) solving social dilemma through concentrated punishment by a leader

研究代表者

大園 博記(OZONO, Hiroki)

鹿児島大学・法文学部・准教授

研究者番号：50709467

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：集団における協力関係が、特定個人への支援集中とその人物による(非協力者および非支援者への)罰行使によって達成されるかを実験室実験により検討した。その結果、例え特定個人に支援集中させた方が集団全体にとって利益が増すと考えられる状況でも、特定個人への強い支援集中はほとんど観察されなかった。これは、権力を委譲することへの忌避を意味しており、本研究は中央集権的な統治に至る抑制要因を明らかにした。

研究成果の概要(英文)：We studied whether and how group cooperation would be achieved by concentration of support for one individual and by the punishment executed by the individual to non-cooperators and non-supporters in laboratory experiments. As a result, we rarely observed the concentration of support even if participants were in the situation where concentration will lead their group to more profits. These results mean that people tend to reject to transferring their power. We suggested that the rejection is one of the inhibiting factor to centralization of human society.

研究分野：社会心理学

キーワード：社会的ジレンマ 罰 制度 リーダーシップ 協力 公共財問題 集権

## 1. 研究開始当初の背景

秩序問題や環境問題等の集団を巡る問題の背景には、個々人による利益追求が集団全体の利益を損なう「社会的ジレンマ (SD)」が存在するとされ、社会科学の諸分野で盛んに研究が成されている。SD を解決する主要なシステムとして考えられてきたのが非協力者に対する罰であるが、罰のコストを誰が負うかを巡って2次的ジレンマが生じるという問題がある。先行研究においては、非協力者を個人的に罰する性質を人は進化的に獲得しており、それにより2次的ジレンマ問題が解決されるとされてきた。このような「個人罰」仮説は一部の実験結果をよく説明する (Fehr & Gächter, 2002) もの、理論的にも (Boyd et al., 2003)、進化環境に近い小規模集団においても (Guala, 2012)、個人罰は稀にしか生じない、といった批判が寄せられている。また、実際の社会を見ても、個人罰はむしろ社会的に抑制され、公権力による集中的な罰システムこそが大規模な協力を支えているといえる。

では、集中的な罰システムはいかにして生まれたのだろうか? その原初的な形は、王や豪族といったリーダーによる集中的な罰 (リーダー罰) だったと考えられる。そこで本研究では、個人罰に代わる新たな SD 解決の枠組みとして、「リーダー罰」を提案する。そして、リーダー罰が自生する条件を実証的に探ることにより、現在の罰システムに至る道筋を描くことを目的とする。リーダー罰の先行研究に松本・神 (2010) がある。彼らはシミュレーションにより、「集団成員からの支援を元手にリーダーが罰する構造 (支援あり構造)」の下では、2次的ジレンマが解決されうるとを示した。リーダーが「SD に協力しない者 (非協力者)」も「リーダーに支援しない者 (非支援者)」も共に罰する「連動罰」をすることで SD 高協力が達成され、リーダー自身も多くの支援を得ることで利益が増し、適応的となるからである。一方、リーダー罰の実験研究は非常に少ない (Baldassarri & Grossman, 2011; O'Gorman et al., 2009)。しかも、これらの実験での罰コストは、リーダーの手持ち資金からのもので、他成員からの支援を受けたものではない (支援なし構造) ため、罰によってリーダー自身の利益が増すことはなく、理論的には2次的ジレンマの解決にいたらない。そこで、筆者は、「支援あり構造」と「支援なし構造」の下で SD は解決されるかを実験室実験で検討した。その結果、「支援なし構造」では、リーダーはほとんど罰行使せず、SD 高協力は達成されなかった。一方「支援あり構造」では、様々なタイプの罰をするリーダーが現れた。分類すると、SD の非協力者もリーダーへの非支援者も両方罰する「連動罰」タイプ、SD 非協力者のみを罰する「利他罰」タイプ、リーダーへの非支援者のみを罰する「利己罰」タイプ、誰も罰さない「無

罰」タイプである。この中で、SD 高協力が達成されたのは、「連動罰」リーダーの下においてのみであった。この結果は、「支援あり」構造で「連動罰」リーダーが生まれた時に SD は解決されることを明らかにした。筆者は、この研究をさらに発展させ、リーダー罰の成立要件を探る研究を計画していた。

## 2. 研究の目的

リーダー罰についての先行研究は、全て「リーダーによる罰と成員のコスト負担の階層構造」が外生的に与えられていた。しかし、そもそもそのような階層構造はいかにして生まれるのだろうか? すなわち、誰もリーダーがいらない社会から、「支援する者-される者」「罰する者-罰される者」という階層構造は、どのような条件がそろえば生まれるのだろうか? この問題は、階層構造・統治構造の創発を扱うものであり、野心的かつ壮大な問いである。

これまで筆者が行ってきた実験手法を修正すれば、この問いに向き合うことが可能になる。具体的には、「リーダーだけにしか支援できない」という制約と「リーダーだけが罰せられない」という制約をなくし、成員全員が誰に対しても支援可能で、誰に対しても罰が可能な状況を設定する。この設定の下では、各参加者は「個人罰」も「他者への支援」も「支援された上での罰」も潜在的には行える。このような状況で、特定の個人に対する支援の集中と、その個人による連動罰が自生する条件を実証的に探ることが本研究の目的であった。

## 3. 研究の方法

本研究における、全ての実験 (実験 5 以外) の構造は類似しており、実験室でコンピュータを介して 6 人集団の実験を行った。実験では、集団での SD (ステージ 1) 支援 (ステージ 2) 罰 (ステージ 3) という流れを繰り返させた。支援ステージでは、自分を含めたすべてのメンバーのうち、誰に自分の資源を提供するかを決定させた。罰ステージでは、支援ステージで得た資源を元手に、誰をいくら減額するかをそれぞれ決定した。このような基本設定の中で、特定個人に対する支援集中と、その個人による連動罰が自生し、安定的な高協力が達成される条件を探った。以下に、本研究で検討した実験条件と、その検討理由を述べる。

### (1) リーダーにふさわしい人物を選ぶ「人気投票」の導入 (実験 1)

人気投票により、誰がリーダーにふさわしいかを成員間で共有できれば、「誰に支援するか」の共有信念を形成され、他成員の集中も予想でき、安心して集中が起きやすくなるだろうと考えた。また、Baldassarri & Grossman (2011) は、投票で選ばれたリーダーによる罰の下では、SD 協力が達成されやすい

ことを示し、それを罰の正統性が担保されたからだと議論しており、本研究においても投票により選ばれた人物による罰により協力が達成されやすいと予想した。

#### (2) 集中により罰効率が上昇するシステムの導入(実験2)

個人がバラバラに罰するより、罰システムを一極集中させた方が、一般に罰の効率は上昇するだろう。それを実験に導入することで、個々人で行うより、特定個人に集中させるインセンティブが生じるだろう。また、集中によって個人間で罰効率に非対称性が生まれるため、報復リスクが減少し、集中された個人は罰を行いやすくなると思った。

#### (3) 罰効率に個人差(力の違い)がある状況の導入(実験3)

実際の社会は、個々人の力(腕力や知力)が均一ではない。当然、強い者に資源を集中させた方が罰の効率がよいため、集中のインセンティブが生じるだろう。また、力が強い者は報復リスクも少ないため、積極的に罰行使できると考えた。

#### (4) 罰行使の意図表明の導入(実験4)

上記の個人差がある状況においては、「本当に力のある者が適切な罰を行ってくれるか」が不確実であり、集中させるのに躊躇が生じるかもしれない。そこで、支援決定前に個々人に「どのような罰をするつもりか」の意図表明をさせることで、その不確実性が低減し、力の強い(罰効率が良い)特定個人に資源が集まりやすいと考えた。また、その個人は、自らの意図表明にコミットしており、適切な罰行使をしやすいと予想した。

また、上記の(4)に関連して、意図表明そのものに協力促進の機能があるのではないかと考え、実験5を行った。なお、この実験5では、支援ステージはなく、SDステージ罰ステージのみで行い、集団人数は4人であった。

#### (5) 協力意図表明による協力促進効果の検討(実験5)

SDでは、他成員が協力するか否かの不確実性が、協力を妨げる一つの要因となる。この問題を解決するには、SD前にそれぞれの協力意図を表明させることが考えられるが、その表明が信頼できるかもまた不確実である。そこで、SD後に罰があれば、「虚偽の表明をしたら、強く罰せられる」と各成員が予期することにより嘘が付きにくくなり、表明の信頼性が高まる可能性がある。つまり、表明は罰と組み合わせられることで不確実性を低減させるため、結果としてSDでの協力は達成されやすいと予想される。本実験では、SDと罰行使を繰り返させる状況を設定し、SD前に自らの協力意図を表明させる「表明」条件と表明させない「非表明」条件を比較し

た。

#### 4. 研究成果

実験結果を端的に述べると、全ての実験(実験1から実験4)を通して、特定個人への支援集中とその人物における連動罰が生じたのは、1グループのみ(全29グループ中)であり、極めてまれにしかリーダーによる統治構造は自生しないことが分かった。以下にそれぞれの実験結果を簡単にまとめる。

##### 実験1: 投票の導入(6グループ)

投票後に、最も罰を課していた人物に対する集中が起こったグループはあったが、その人物が安定的に連動罰を課し続けることはなく、高協力は形成されなかった。また、非協力者が協力者に支援をする“賄賂”行動が広く観察され、実際に支援をした非協力者を罰する協力者はほとんどいなかった(賄賂が効果的であった)。この賄賂行動は、実験2, 3, 4においても、広く観察された。

##### 実験2: 集中による罰効率の上昇システムの導入(9グループ)

多くのグループで一貫した罰行使者は現れず、集中はおこらなかった。また、一貫した罰行使者がいても、弱い集中(6人中2人程度の支援)しか起こらず、高協力は形成されなかった。

##### 実験3: 力の個人差の導入(11グループ)

1グループにおいて、最強者(もっとも罰効率が良い者)による一貫した連動罰とその人物への支援集中が起きた。これが唯一の統治の自生ケースであった。しかし、他のグループでは、最強者が罰をほとんどしなかったり、したとしても行動に一貫性がなかったり、たとえ一貫性があっても支援が集まらない(最強者が非支援者を罰しなかったため)など、統治行動は自生しなかった。

##### 実験4: 罰行使の意図表明の導入(3グループ)

罰行使を表明しても、実際にはしない者が初期から多数現れ、表明は形骸化し高協力は達成されなかった。

以上のように、支援集中を促進すると考えられる様々な要因を導入した下で実験をおこなったが、1グループを除いて、特定個人への支援集中とその人物による連動罰は自生しなかった。支援する側とされる側の双方の視点から、その理由を考察する。まず、支援する側には、特定個人に権力を委譲し、集中させることへの忌避が強く見られた。集団のためになりそうな個人が現れても、強い集中が容易には起きなかったのである。これまでの研究から、実際に権力を得た個人が集団のためにはならない行動を選択する可能性があることは示されている(大園ほか、2012;

Nosenzo et al., 2014)。このことから、権力者への不信や警戒感を示すのは適切な反応と言えるが、それが統治による高協力を阻害する可能性を本研究は示している。近年、リーダーや集権制度によって、秩序維持は促進されるかを問う研究は増加しており (Sutter et al., 2014; Traulsen et al., 2012))、それには権力への不信を乗り越える仕組みが必要であることを本研究は示している。

一方で支援される側には、自らが特別な立場につくことへの忌避が見られた。自らへの支援集中が起こると、罰行使をしなくなるケースが多く観察されたのである。また、そもそも支援集中を得るだけの一貫した行動をしない被験者も多くみられた。これらは、連動罰を課しても自らの利益が高まるという長期的な見通しが持てないことが原因として考えられる。また、支配的な立場に立つことで、恨みを買う恐れを感じ、そのような立場につくのを拒絶したのかもしれない(本実験状況においては、報復が可能な状況であったため)。

以上のように、支援する側、される側双方に「特定個人による統治」を拒む傾向があることが示唆された。しかし、実際の社会では、このような統治は、歴史上何度も起きてきたのは明らかである。今回の実験室と歴史の乖離はどう説明できるのであろうか。実験室が実際の社会と大きく異なるのは、個々人の特性が暴露されにくいという点である。実際の社会では、様々な相互作用から個々の人間の能力や人格を推し量ることができ、「誰がリーダーにふさわしいか」を集団内で共有するのが容易である。そのため、特定個人への支援集中は起こりやすいだろう。しかし、匿名性が保持され、相互作用が限定的な実験状況では、そのような共有は困難であり、不確実性の高さから、特定個人への集中を忌避する傾向が観察されたと考えられる。今後は、より互いの特性が明らかになる実験状況を設定し、いかなる条件で特定個人による統治が自生しやすいのかをさらに検討していく必要がある。

また、実験5においては、協力意図表明による協力促進効果の検討をした。その結果、予測に反し、表明条件と非表明条件でSD協力率の違いは初期から見られなかった。さらに、初期の罰行動を分析すると、表明条件の方が非表明条件より非協力者への罰行使が少ないことがわかった。この罰の抑制は、表明が虚偽であるか否かによらなかった。つまり、「協力するか否か」の意図表明があるだけで、罰行使が抑制されることがわかった。本実験では虚偽の表明が初期から蔓延していたため、協力を目指す動機づけが弱まり、罰行使が抑制されたのかもしれない。これまでの研究においては、虚偽の表明は強い罰行使を招くことが示されており (Brandts & Charness, 2003; Ohtsubo et al., 2010; 大

園ほか, 2013)、本実験はそれらの結果とは矛盾しており、興味深い。先行研究は一回限りの取引であったが、本実験は繰り返しがあつた。そのため、長期的な協力が見込めないことから来る諦めが生じやすかつたのかもかもしれない。今後、関係の継続性が虚偽表明とそれに対する罰行使に及ぼす影響を検討することで、協力形成に表明やコミュニケーションが果たす役割を明らかにしていきたい。

#### (1) 実験

##### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 2 件)

大園博記, 社会的ジレンマでの罰行使における協力意図表明の影響, 日本人間行動進化学会第7回大会, 2014年11月29日, 神戸大学(兵庫県神戸市)。

大園博記, 神信人, 渡部幹, 清水和巳, 社会的ジレンマ状況で集権的罰は自生するか?: 実験研究, 日本人間行動進化学会第6回大会, 2013年12月7日, 広島修道大学(広島県広島市)。

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

特になし

#### 6. 研究組織

##### (1) 研究代表者

大園 博記 (OZONO, Hiroki)

鹿児島大学・法文学部・准教授

研究者番号: 50709467