

社会規範と利他性

—その発現形態について—

飯 田 高

I. はじめに

自生的な秩序形成がいかにして可能なのかという問題は、いまや社会科学のみならず、「自然科学系」とされる諸分野でもさかんに論じられている。たとえば、進化生物学、生態学、現象数理学、社会物理学といった分野では、社会秩序の形成過程をモデル化して解析する試みが数多く行われ、それぞれに興味深い知見を提供してきた¹⁾。

複数の個人²⁾が存在する場面では、よほど特殊な環境でない限り、個人間の相互作用が生じる（ここでの相互作用には、対面のコミュニケーションや取引をはじめとする直接的な相互作用だけでなく、第三者を介した間接的な相互作用も含まれる）。このとき、いつの間にか秩序が出来上がり、その秩序に人々が拘束され、さらに秩序が再生産される、ということがしばしば起きる。「社会規範」は、このような自生的な秩序形成（あるいは秩序形成装置）を表す言葉として用いられることが多い。

この意味での社会規範をどのように概念化するかをめぐっては、大きく分けて2つの立場が存在する。ひとつは、他者の行動に関する期待が人々の間に成立し、それに従う行動が自己拘束的となっている状態をもって社会規範が生成されたと考える立場である (Lewis 1969; Young 2008)。言い換えれば、人々の間で相互理解ができ、他者もそれに従うだろうという期待のもとで、自分もまたその相互理解のとおりに行動することを選好する、と

1) 書籍の形をとっているものとしては、たとえば Sigmund (2010), Bourke (2011), Nowak (2011), Nowak and Coakley (2013) など。これらの分野の内容を知るうえで非常に有用な文献として、中丸 (2011) を挙げておく。

2) 進化生物学や生態学の研究では、人間以外の他の生物を対象としていることがある。もちろんその場合は「個人」ではなく「個体」だが（英語ではどちらも“individual”）、本稿が扱うのはほとんど人間の社会なので、利他性の進化を扱う II 3 以外の箇所ではすべて「個人」と表記している。

いう状態である。

もうひとつは、人々の「協力行動」をもたらすルール、すなわち、他者にとって利益になる行為を促進したり、逆に他者にとって有害になりうる行為を抑制したりするルールが確立している状態をもって社会規範が生成されたとみなす立場である (Bicchieri 2006)。自己拘束性があるという点では前者と同じだが、協力行動と関係するか否かを前者が問題としていない分、後者のほうが範囲はより狭くなっている³⁾。以下では、後者のような社会規範を特に「協力規範」と呼ぶことにしたい。

本稿の目的は、社会規範の研究において「利他性」の概念がどう用いられているかを確認したうえで、その「利他性」が社会規範にどのように反映されうるかについて、従来とは異なる観点から考察を加えることにある。

II以降の構成は以下のとおりである。IIでは、諸分野における協力規範の研究をレビューし、協力規範の構成要素とされる「利他性」を検討する。次いで、「利他性」の存在を説明する理論を概観する。IIIでは「利他性」がどのような現象——望ましくないと思われる現象——を生み出しうるかについて述べる。IVでは、「利他性」が社会規範といかなる関係に立つのかを再考する。そこで強調されるのは、「社会規範がどのように利用されるか」という視点である。

II. 社会規範を成り立たせる要素

1. ジレンマ・パラダイム

現在の協力規範研究できわめて大きな影響力をもっているのは、「囚人のジレンマ」または「社会的ジレンマ」のモデルである。むしろ、これらのジレンマ・モデルが協力規範の研究の原動力となってきた、と言ってよい。

ジレンマ・モデルが表現しているのは、個人にとっての利益と、他者を含めた「社会」ないし「集団」全体の利益が衝突している状況である。個人は自分の利益を最大化する行動を選択しようとするため、そのまま放っておくと全体の利益は損なわれる。つまり、ジ

3) 前者に含まれて後者に含まれないルールの例としては、道路のどちら側を通行するかを指示するルール、特定の業界における取引の方法に関するルール、適切な敬語についてのルールなどが挙げられる。なお、カウシク・バス (Kaushik Basu) は、社会規範をその役割の観点から「合理性制限規範」(rationality-limiting norm)、「選好形成規範」(preference-changing norm)、そして「均衡選択規範」(equilibrium-selection norm)の3種類に分けている (Basu 1998:477-8)。大まかに言えば、前者の意味での(自己拘束的なルールとしての)社会規範は均衡選択規範、後者の意味での(協力行動を導くルールとしての)社会規範は合理性制限規範に相当する。

レンマ・モデルにおいては均衡状態と全体にとって望ましい状態とが乖離しているの、全体の利益を最大化するためには各人の自己利益追求を何らかの手段で制約する必要がある。

一般に、社会規範はそのような手段のひとつと考えられている。たとえば、環境保護や経済取引に関する規範、共同作業における任務の懈怠を防ぐ規範、共有資源の過剰利用を控えさせる規範などは、いずれも自己利益追求に歯止めをかけて全体の利益を増進する内容となっており、本稿で言う「協力規範」の具体例となっている (Ostrom 1990)。

したがって、ジレンマ・モデルに立脚した社会規範研究は、「いかにして均衡状態を移動させ、全体にとって望ましい状態——効率的な状態——を実現するか」という問題に取り組むことになる。これは、協力規範を実効化 (エンフォース) するメカニズムを問うことに等しい。この問いに対して、個人の合理的な選択を考察の出発点とする立場からは、概ね次のような回答が示されてきた⁴⁾。

(1) 繰り返しゲームにおける実効化 (直接互惠性)⁵⁾

囚人のジレンマを繰り返しプレーする設定になっているとき、一定の条件のもとで、各プレーヤーが協力行動を選択するという状態が均衡となりうる (フォーク定理)⁶⁾。各プレーヤーは相手が裏切った際に「これからは協力しない」という形でサンクションを与えることができるので、相互協力の状態が自己拘束的になりうる。

繰り返しゲームによる説明の主な難点は、社会や集団が大規模な場合、相互協力が維持される条件が非常に厳しくなる、という点である (Bowles and Gintis 2011: 63-67)。「誰かが裏切ると自分も将来の協力を差し控える」という方針を各人がとっているならば、一部の人たちが裏切り行動をとるだけで相互協力の状態はたやすく崩れてしまう。

理論的な難点にもかかわらず、繰り返しゲームによる説明は直観に訴える力が大きく、上記の問いに対する標準的な回答でもある。現実の社会でも「あなたが助けてくれれば、お返しをしてあげよう」あるいは「あなたが裏切れば、もう協力しない」という行動指針はたしかに存在し、世界中で広く見られるいわゆる「黄金律 (Golden rule)」に反映され

4) これは網羅的なリストではなく、紙幅の関係で一部の主要なモデルに限定してある。なお、簡単なモデルの分類については飯田 (2004: 68-69) も参照 (ただしそこでは、協力規範に限らず社会規範一般を分類している)。

5) たとえば, Trivers (1971), Axelrod (1984), Taylor (1987), Ellickson (1991) などが繰り返しゲームを用いて協力行動の説明を試みている。

6) 十分に割引因子の高い (割引率の低い) プレーヤー同士が無限回の繰り返しゲームを行うとき、協力行動を各人がとる状態が均衡となりうる。有限回であっても、弱い均衡概念を採用すれば協力行動が均衡として生じうる。繰り返しゲームやフォーク定理についての詳細は、岡田 (2011: 231-267), 神取 (2015) を参照。

ている (Neusner and Chilton 2008).

このような行動指針は、下記の間接互惠性との対比で「直接互惠性 (direct reciprocity)」と呼ばれることがある。

(2) 評判を通じた実効化 (間接互惠性)⁷⁾

人間社会では、第三者の目と口が個人の意味決定に強い影響を及ぼすことがある。

たとえば、AがBを裏切ったとする。その情報が第三者のCに伝わると、Cは「Aと交流するのは避けよう」と思ったり、「Aに対しては自分も裏切っても構わないだろう」と考えたりする可能性がある。逆に、AがBに対して過去に協力行動をとっていれば、それを聞いたCが「自分もAに対して協力行動をとろう」と思うかもしれない。

このようなメカニズムで人々の相互協力が達成されるためには、プレーヤーの性質や過去の行動に関する情報、すなわち「評判 (reputation)」が人々の間に流通していることが必要である。協力行動をとらない個人は社会や集団の中での評判を低下させ、交流相手を見つけて利益を得ることが難しくなる。したがって、もし評判の情報が個人の性質を識別する指標としてうまく機能しているのであれば、自分自身の利益を最大化しようと思う個人は協力行動をとるようになる⁸⁾。

評判をメンバー間で共有することによって達成されるこの種の互惠性は「間接互惠性 (indirect reciprocity)」と呼ばれる場合がある (Alexander 1987)。

(3) 別種のメカニズムによる実効化

合理的選択に基づくモデルで通常仮定されているのは、サンクションの行使にコストがかかるということである。コストがかかる場合は、そのコストを凌駕する利益が得られない限り誰も逸脱者に対してサンクションを科そうとはしないだろう。だが、もしサンクションの行使にコストがかからなかったり、何らかの理由によってそのコストが埋め合わされたりするのであれば、より頻繁かつ広範にサンクションが行使され、協力行動も多くなるはずである。

コストの埋め合わせの例をいくつか挙げよう。まず、わざわざサンクションを皆のために行使した人は、他者から高い評価を受ける可能性がある (McAdams 1997)⁹⁾。人々は

7) Sugden (1986), Kandori (1992), Brandt and Sigmund (2004), Ohtsuki and Iwasa (2004)などを参照。

8) 過去の行動の履歴がどのように評判に反映されるか、評判をもとにどのような戦略を選択すれば協力状態が導かれるかといった問題については、Ohtsuki and Iwasa (2006), 巖佐 (2015), 真島 (2015)などを参照。

9) リチャード・マカダムズは、「規範を形成する根源的な力は、尊敬や威信への欲求、つまり、他者からの相対的な評価への欲求」だと述べている (McAdams 1997: 346)。

他者と比較した相対的な評価を気にすると仮定し、さらに、他者を称賛する（プラスの評価をする）のにはコストがかからないと仮定する。この場合、もし他者からの評価が個人の効用関数に組み込まれているとすれば、一定の条件が揃ったときにサンクションが行使されうるだろう¹⁰⁾。

また、サンクションを行使することが協力的な性質の現れだとみなされることもあるかもしれない (Posner 2000)。この説明によると、サンクションの行使は信憑性のあるシグナルとして作用する。つまり、良いタイプの人（衝動的でない人）だけが負担できるようなコストを自ら負うことができれば、自分の性質を知らせる効果的なシグナルになるだろう。逸脱者に対してサンクションを科すという行動は、そのようなコストの一種と位置づけられる。

以上のモデルは、個人の合理性の仮定をできるだけ維持したまま協力規範を説明しようとしている。これらはいずれも囚人のジレンマや社会的ジレンマをベースにしており、協力規範の実効化メカニズムをプレイヤーの「自己利益」（この中身は論者により異なる）の点から解明することを試みている。

2. アノマリー

しかし、個人の完全な合理性のみでは協力規範の生成や維持を説明しにくい。そこで、個人の合理性についての条件を緩和したり、個人の選好や認知についての条件を加えたり、あるいは個人が直面している環境に特殊な条件を課したりする協力規範モデルが提示されてきている¹¹⁾。

このような動きと並行して、実験の手法による研究や、協力規範の調査研究が積み重ねられてきた (Ostrom 1990; 2000)。なかでも行動経済学における実験研究は、実際の人々の行動が理論上の予測とどのくらい乖離しているかを少しずつ明らかにしている。これらの実験研究で確認されている現象のうち、協力規範と関連が深いのは「理論の予測を超える協力行動」と「非協力者に対する処罰行動」であろう。以下、それぞれに関する主な実験研究を紹介する¹²⁾。

10) この点については、Brennan and Pettit (2004) の議論も参考になる。

11) たとえば、松井 (2002)、Akerlof and Kranton (2005; 2010) などを参照。McAdams (1997) のモデルも、個人の選好に関する特別な仮定を含んだモデルのひとつである。

12) 行動経済学ではほぼ必ず参照される著名な実験研究のみを紹介している。すでにこのような実験研究をご存知の方は、本項を読み飛ばして3に進んでいただきたい。

(1) 理論上の予測を超える協力行動

囚人のジレンマや社会的ジレンマの状況を実際の人間がプレーすると、経済理論の予測を上回る協力行動が観察される。

社会的ジレンマ状況での人間行動を調べる際、用いられる頻度が最も高いのは「公共財ゲーム」の実験である。この実験においては、すべての実験参加者（プレーヤー）が資金の拠出をすれば全体の利益は最大になるが、利益はグループ全体に分散されるため、個々のプレーヤーにとって合理的な選択は「拠出しない」という行動になる¹³⁾。

たとえば、4人のプレーヤーからなるグループを作り、実験者は最初にプレーヤー1人につき1,000円の元手を渡しておくとする。各プレーヤーは、1,000円のうちいくらかをグループ全体のために拠出するかを決める。実験者は、集まった拠出額の合計を3倍にしてグループに返還する。そして、返還された金額を均等割りした額が各プレーヤーの利得となる。仮に全員が1,000円全額を拠出したとすると、4人から集まった4,000円が12,000円に増えるので、最終的には全員が3,000円ずつもらえる。このとき、全員が得る額の合計は最大となる。

ここで、4人のうちの1人が拠出額をゼロにしたとしよう。この場合は集まった3,000円が9,000円になり、1人あたり2,250円が返還される。拠出しなかったプレーヤーの最終的な取り分は、手元に残っている1,000円と合わせて3,250円となる。

この例からわかるとおり、各プレーヤーは資金を拠出しないインセンティブをもつはずである。だが、公共財ゲームの実験によれば、プレーヤーの行動にはかなりのばらつきがあり、元手の40～60パーセントを拠出するという行動が多い(Ostrom 2000: 140)。まったく拠出しないという人もいるが、決して多数派ではない。

(2) 相手方による処罰行動

経済学者のエルンスト・フェア (Ernst Fehr) とシモン・ゲヒター (Simon Gächter) は、処罰オプション付きの公共財ゲームを用いた実験を実施し、サンクションの可能性が協力行動にいかなる影響を及ぼすかを調べている (Fehr and Gächter 2002)。この実験では、プレーヤーが公共財ゲームを複数回行い、ゲームの最後に非協力者に対して処罰ができるようになっていた。ただし、処罰を行うと処罰者自身にもコストがかかり、利得はその分

13) 公共財ゲームの内容を形式的に表すと次のとおり。n人のプレーヤーからなるグループがあり、各プレーヤーは公共財のための拠出額を決定する（プレーヤー*i*の拠出額を c_i とする）。実験者は拠出額の合計を*t*倍し（*t*は $1 < t < n$ を満たす数）、等分したうえで各プレーヤーに分配する。このとき、プレーヤー*i*の利得は $t \sum_{m=1}^n c_m / n - c_i$ となるので、拠出額をゼロにするのが個人にとって合理的な選択となることがわかる。

だけ減少する（なお、誰が処罰したかは他の人からはわからないようになっていた。したがって、処罰行動が処罰者の評判を高めることはない）。

結果だけ述べると、処罰オプションがある場合は協力率が大きく上昇し、しかもそれが最後のラウンドまで高い水準で維持された。そして、実験参加者はコストのかかる処罰行動をとるのを厭わなかったのである¹⁴⁾。

もしかすると、相手が次のラウンド以降に協力行動をとるよう促すべく、警告の意味で処罰したのではないか、と思われるかもしれない。そうだとすれば、処罰者は自らの将来の利得を増やそうとして罰を行使しているのであるから、処罰行動は依然として自己利益に基づいた行動だと言える余地がある。

そこでフェアとゲヒターは、グループのメンバーがラウンドごとに入れ替わるという条件のもとで、処罰オプションつき公共財ゲームの実験を行った。この条件のもとでは、処罰行動は自分自身の利益につながらない。せいぜい被処罰者と将来関わる人を利するくらいである。ところが、この条件でもやはり処罰行動は観察された。

たとえ処罰者の利益に直結しなくても、それどころか自分にとってはコストにしかなくとも、被処罰者と関わる可能性のある人たちのために処罰を行うのかもしれない。フェアとゲヒターは、このような処罰行動を「利他的処罰 (altruistic punishment)」と名づけている。

(3) 第三者による処罰行動

利他的処罰の話にはさらに続きがある。この議論をもう少し突き進めると、当事者のみならず第三者も処罰行動をとると予測されるだろう。

たとえば、ゲームの経過を第三者 Z が観察しており、Z はプレーヤー X に対して罰を与えることができるとする。具体的に言うと、Z は一定額の金銭を X の取り分から減らすことができるが、先の処罰オプションつき公共財ゲームと同様、罰を与える際には Z 自身もコストを負担しなければならない。

実際にこのゲームを行ってみると、たとえコストがかかっても Z は処罰行動を選択することがある。先に触れたエルンスト・フェアとウルス・フィッシュバッシャー (Urs Fischbacher) は、囚人のジレンマのゲームを使ってこの傾向を確認している (Fehr and Fischbacher 2004)。

14) ちなみに、処罰のコストが高ければ処罰は減少する (Anderson and Putterman 2006)。したがって、自己利益の追求という要素も処罰には絡んでおり、おそらく感情と計算が同時に働いているのであろう。

囚人のジレンマでは、直感的にも理解されるように、相手が協力行動をとっているにもかかわらず非協力行動を選択したプレイヤーが処罰の対象とされやすい（つまり、一方的に得をした人を罰する）¹⁵⁾。だが、両者が非協力行動を選択した場合も、その率は低くなるものの、第三者は双方のプレイヤーに対して罰を与える傾向がある。つまり、単純にプレイヤー間の公平性の観点のみから処罰を加えるわけではないということを示しており、この点は注目に値する。

3. 利他性

(1) 協力規範を支える利他性

前項2で述べたアノマリーの背後には、「利他性 (altruism)」が作用していると考えられている。自己の利益は犠牲になるが他者の利益には資する行動を利他行動 (altruistic behavior)、そのような行動を選択させる性質がここでの利他性である（「利益」とは何かが問題になりうるが、ここでは物質的な利益と精神的な利益の双方を含むものとしよう）¹⁶⁾。

たとえば、第三者による処罰行動は処罰者自身の利益には何ら寄与せず、他の人たちにとって得になりうるというだけである。また、公共財ゲームにおける協力行動も、フリーライドの機会を自ら手放して皆の利益に貢献しているので、ここで言う利他性の現れと解釈することができる¹⁷⁾。

前項で触れた実験結果を参照すると、協力規範が維持されるためには利他性の要素がどこかで必要であることが示唆される。翻って私たちの日常生活を考えてみても、人間の利他性を示すと思われるエピソードはそこかしこに転がっている。たとえば、自然災害の被災者に救援物資を送ったり被災地でボランティア活動を行ったりする、慈善団体や学校に匿名の寄付をする、見つけた落とし物を交番に届ける、駅のホームから線路に落ちた酔っ払いを助ける、といった行為がその例である。

この例で見られるような利他性は、社会化の過程で身につけられる部分もあるが¹⁸⁾、す

15) フェアとフィッシュバッチャーの実験によると、協力した相手に対して裏切ったプレイヤーに対しては、半程度程度の第三者 (Z) が処罰を行っている (Fehr and Fischbacher 2004: 83)。

16) 本稿では、「利他性」や「利他的」という語を、ある人間が自分自身の利益を犠牲にして相手の利益にかなうことをする場合に限って用いる。要するに、自分にとってはマイナス、相手にとってはプラスとなる場合である。「利他性」の語をこのような用法に限定している例として、Bowles (2004:110) を参照。

17) 利他性の現れと解釈される行動は、社会的ジレンマ型の実験で観察される行動だけではない。「最後通牒ゲーム (ultimatum game)」において低い金額をオファーされた回答者が拒否する、という行動も利他性の例として言及されることがある。

18) 上記の利他行動がどの程度とられるかは、明らかに文化や制度も影響している。たとえば、拾得物を交番に

べてがそうだというわけではなさそうである。

認知心理学者マイケル・トマセロ (Michael Tomasello) らの研究グループは、小さな子供を対象とした実験を何種類もやっている。一般に、生後18ヵ月ごろの幼児は、特定の社会の価値観を身につける前の発達段階にある。そのような幼児であっても困っている人を見ると手助けをしてくれる、というのがトマセロらの観察結果である (Tomasello 2009; Warneken and Tomasello 2009)。

たとえば、成人である実験者が洗濯ばさみを誤って床に落としたとする。または、手のふさがっている実験者がキャビネットの扉を開けようとしているとする。このような場面に居合わせた幼児は、相当高い確率で救いの手を差し伸べてくれる¹⁹⁾。つまり、他者が何をしようとしているかを理解したうえで手助けをしてくれるのである。このとき、幼児は自発的に援助行動をとっており、外から与えられる報酬はかえって逆効果となることが知られている。

トマセロらの実験は、利他性が生得的に備わっている可能性がある、ということを示すものと言えよう。それと同時に、他者が何を欲しているかを認識する能力が幼いうちにある程度出来上がる、ということを示す研究でもある。

(2) 利他性の進化

利他性が生得的に備わっているとすると、生物進化の文脈で大きな問題に直面することになる。利他性はどのようにして進化しうるのか、という問題である。

生物にとっての利益とは、端的に言えば「子孫を多く残す」ことである。ある生物個体がどれだけの子孫を残せるかを数値で表した尺度を「適応度 (fitness)」と呼ぶが、この表現を使うと、生物にとっての利益は適応度であり、利他行動とは「自らの適応度を低下させ、かつ、他の個体の適応度を上昇させる」行動だということになる。

利他行動をとる個体は、非利他的な行動をとる個体と比較して不利である。グループ内に利他的な個体と非利他的な個体が存在しているとすると、前者は相対的に適応度が低くなる(=子孫が相対的に少なくなる)ため、そのグループからは利他的な個体は次第に淘汰されていく。つまり、自然選択(自然淘汰)によって利他性はやがて消滅してしまうのである。

届けるか否かについての意思決定は、制度の違いによって大きく変わりうるだろう (West 2005)。だがその一方で、文化や制度の違いを超えた利他性もまた存在し、自然科学の協力規範研究はそのような意味での利他性を念頭に置いていると言えよう。

19) 人間に近い動物種であるチンパンジーも、ある程度の援助行動はとってくれるという。しかしながら、人間ほど広い範囲の場面で援助行動を見せてくれるわけではない (Warneken and Tomasello 2006)。

このような利他性がいかにして逆境に抗して生き延びたかを説明する主要な理論としては、①血縁選択、②互惠的利他主義（直接互惠性）、③マルチレベル選択が挙げられる²⁰⁾。

①血縁選択

個体自身の子孫ではなく、個体の遺伝子を残すことをその個体の利益と捉えれば、利他性が進化しうることを示せる。進化生物学者のウィリアム・ドナルド・ハミルトン(William Donald Hamilton)は、「血縁度」および「包括適応度 (inclusive fitness)」の概念を用いて利他性の進化する条件を論じている (Hamilton 1964)。

ハミルトンの論文や用語の定義は難解であるが、中心となっている命題自体は、「行為者と受益者の間の血縁度」と「利他行動の利益」の積が「利他行動のコスト」を補ってもなお余りあるとき、当該行動は自然選択を通じて進化しうる、というシンプルなものである²¹⁾。このような血縁者の中での利他性の進化を「血縁選択 (kin selection)」と言う²²⁾。

②互惠的利他主義

ただ、利他行動は血縁関係にある者同士だけに見られるわけではない。血縁関係のない者の間で観察される利他性を説明するうえで有効な理論とされるのが、直接互惠性ないし「互惠的利他主義 (reciprocal altruism)」の理論である。

互惠的利他主義を明確な形で示したのはロバート・トリヴァーズ(Robert Trivers)であった (Trivers 1971)。利他行動をとった個体が大きなコストを一時的に負担したとしても、将来自分が困った状況に陥った場合に受益者から利益を得ることができれば、結局は自分にとって得になる。このような行動は短期的には個体にとってコストになっているものの、長期的に見れば利益になるため、自然選択で生き残ることができる²³⁾。

20) 詳細は Nowak (2006), Bowles and Gintis (2011) などを参照。II 1 (2) で言及した間接互惠性も、利他性の進化を説明するモデルとして挙げられる (ただし本稿では、合理性を基礎とした協力規範のモデルとして紹介した)。

21) たとえば、「自分のきょうだいの子を助ける」という援助行動によって「きょうだいの子」(受益者) が 10 の利益を得るとする。このケースでは行為者と受益者の間の血縁度は 1/4 だから、コストが 2.5 未満である限り、行為者にとっても最終的に「利益」となる。ここでの「利益」とは、行為者の包括適応度が上昇するということである。これは「ハミルトンの法則」と呼ばれる (ただし、ハミルトンの法則は血縁選択だけを対象としているわけではなく、実際にはもっと射程の広い法則である)。

22) 包括適応度の概念は、個体の利益を「個体自身の子孫を残すこと」ではなく「個体の遺伝子を残すこと」へと転換するものであった。リチャード・ドーキンス (Richard Dawkins) の「利己的な遺伝子」はその考え方をさらに一歩進め、個体の利益ではなく遺伝子の利益を中心に据える構図を描いたものとして位置づけられる (Dawkins 1976)。

23) この理論は、本質的には II 1 (1) の繰り返しゲームと同じであり、トリヴァーズ自身もそう述べている。ただし、

互惠的利他主義を遺伝子のレベルで考えたとき、血縁選択と互惠的利他主義の大きな違いは次の点にある。すなわち、血縁選択において得をする（生き残る確率が高くなる）のは受益者の遺伝子のコピーであるのに対し、互惠的利他主義において得をするのは行為者自身の遺伝子のコピーだという点である。血縁選択の場合、行為者と受益者が遺伝子を共有しているため行為者の遺伝子も結果的には得をする。他方、互惠的利他主義の場合には受益者を介することによって行為者の遺伝子が得をする、という仕組みになっている²⁴⁾。

トリヴァーズによると、人間の心理（特に感情システム）はこのような利他行動を動機づけるように進化した（Trivers 1985）。たとえば、裏切った人に対して憤激して攻撃したり、他者を騙した場合に罪悪感を抱いたりするのは、互惠的利他主義の基礎をなす感情の一例である。

③マルチレベル選択

哲学者エリオット・ソーバー（Elliott Sober）と生物学者デイヴィッド・スローン・ウィルソン（David Sloan Wilson）は、上記のメカニズムを包摂する見方を提示している。彼らは、個体レベルで働く自然選択以外にも集団レベルで働く自然選択もあり、さまざまなレベルでの自然選択を通じて利他性が生き残りうる、と論じる（Sober and Wilson 1998）。

たとえば、閉鎖的で小規模な集団が複数存在しているとすると、集団間の交流は行われていない反面、集団内の交流は活発に行われているため、それぞれの集団内ではメンバーの性質が似通っている、と仮定する。さらに、集団はお互いに隔離されていて、集団間の差異は大きくなっているとしよう。

集団内部では、自然選択の力によって、利他行動をとる個体は次第に減っていくかもしれない。ところが、利他的な個体が多い集団はそうでない集団よりも（定義上）生き残りやすくなるはずである。利他的でない個体が多い集団は徐々に駆逐されていくだろう。一方、利他的な個体が増殖して一定数に達すれば、利他的な個体同士で新たな集団を作ることできる。そのようにできた集団は、他の集団よりも生き残る確率が高くなると考えられる。

これは、集団レベルで自然淘汰が起こるとする「集団選択（group selection）」の理論

プレイヤーの合理性や予見能力をどのように仮定しているかは異なっているので、その点は注意が必要である。トリヴァーズの理論の舞台になっているのは生物進化のプロセスであり、そこではプレイヤーの合理性は仮定されていない。

24) したがって、行為者と受益者が遺伝子を共有している必要はなく、異なる生物種の間でも互惠的利他主義は成り立ちうる。

を新たな観点から再生したものと言えるが、複数のレベルでの選択があることを強調する意味で、ソーバーとウィルソンは「集団選択」ではなく「マルチレベル選択 (multilevel selection)」と呼んでいる²⁵⁾。

マルチレベル選択の理論は近年有力になりつつあり、たとえばロバート・ボイド (Robert Boyd) らは利他的処罰の存在をマルチレベル選択の観点から説明している (Boyd et al. 2003)。マルチレベル選択の理論を支持する論者の説明によると、一般の生物種の場合は集団レベルでの淘汰圧は個体レベルの淘汰圧と比べて弱いが、人間社会では文化や制度が確立されているために集団間の競争が激しくなる。そのため、集団レベルでの淘汰圧が無視できないほどの力を持ち、集団を益する利他性が進化する可能性は高くなる²⁶⁾。

III. 利他性のダークサイド

1. 偏狭な利他性

今までの議論を振り返っておこう。社会規範、ことに協力規範がいかにして存在しうるかは、ジレンマ・モデルに基づいて説明されることが多い。協力規範を個人の合理性から説明するのも一応可能であるが、協力規範が維持されるための条件が厳しくなるうえ (II 1)、実際の人間行動や種々の実験結果を見ると、別種の心理メカニズムを通じて協力規範が維持されていることが示唆される (II 2)。ここでよく持ち出されるのが「利他性」であり、なぜ人間が利他行動の基盤となりうる心理メカニズムを備えるようになったかを説明する理論もいくつか存在する (II 3)。

25) 集団選択 (群選択または群淘汰とも呼ばれる) の可能性についてはダーウィンも示唆していたが、集団選択の効果は個体選択の効果と比べて非常に弱いものでしかないことがウィリアムズによって指摘されて以降 (Williams 1966)、集団選択の理論は力を失っていた。マルチレベル選択の理論は、この集団選択の考え方を「形質集団」の概念を用いて再定式化した理論だと言うことができる。

26) まず、 n 個の集団が存在し、各集団を構成する個体の数を z_i ($i=1,2,\dots,n$) とする。そして、個体が残す子孫の数 (適応度) は集団ごとに異なり、一個体あたり w_i ($i=1,2,\dots,n$) の子孫を残せると仮定する。そのとき、集団内の個体数の平均値は $\bar{z}=(1/n) \sum z_i$ で表されるが、次の期にこの数値がどのように変化するか (これを $\Delta \bar{z}$ で表す) を教えてくれるのが「プライス方程式」である。これは

$$\Delta \bar{z} = \text{cov}(z_i, w_i) / \bar{w} + E(w_i \Delta z_i) / \bar{w}$$

という式で表される ($\text{cov}(z_i, w_i)$ は z_i と w_i の共分散、 \bar{w} は w_i の平均値 (母集団全体の平均適応度)、 $E(w_i \Delta z_i)$ は一個体の子孫数がどれくらい変化するかを示す期待値である)。

この式から導き出されるのは、ある遺伝子が増加するか否かは「集団内で起こる個体レベルでの自然選択」と「集団間で起こる集団レベルでの自然選択」の2つの効果に依存する、ということである。そして、後者 (集団レベルでの自然選択) の効果は、遺伝子の割合と集団規模の共分散の大きさによって決まる。また、前者の効果が小さければ、集団レベルでの自然選択の影響が増す。

たしかに、利他性は人間の社会を支える大事な要素なのかもしれない。しかし、利他性は望ましい社会状態をいつでも実現してくれるわけではない。利他性はかえって有害な結果を生じさせることがある。

社会規範との関係で最も問題になりやすいのは、利他性の「他」がいったいいかなる範囲の人々を含んでいるのか、という点である。II 3 で紹介した利他性の進化の理論からは、利他性は一定の範囲、それも比較的狭い範囲の人たちしか含まない、ということが推測される。血縁選択とマルチレベル選択については言うに及ばず、互惠的利他主義の理論でも、利他行動が進化するための条件が具備されやすいのは小規模の社会や集団である。

もし利他性が周囲の人々にだけ向けられるもので、しかも、心理メカニズムという間接的な装置によってのみ発動されるものなのだとすれば、手放して称賛されるものではなく、それどころか全体としては大きな損失を引き起こしかねない。

実際、人間はグループというものに敏感であり、自分が所属しているグループ（内集団）とその他のグループ（外集団）とで扱いを区別することが多々ある²⁷⁾。なかでも、内集団の人たちを外集団の人たちよりも好意的に見たり優遇したりする（内集団びいき・身内びいき）のは、日常生活でも実験でも観察されるところである²⁸⁾。

さらには、内集団の人たちに対しては配慮を示す一方で、外集団の人たちを同じ種類の人間として扱わない「非人間化 (dehumanization)」と呼ばれる現象も起こりうる (Haslam 2006)。人間を人間として見ないという場合もあれば（奴隷社会が例）、外集団に属する人を嫌悪や軽蔑の対象として見るという場合もある（緊張関係にある国家が例）。一般的に言って、内集団ネットワークの強化は外集団の非人間化を助長する傾向がある (Waytz and Epley 2012)。

マルチレベル選択の理論に従えば、このような「偏狭な利他性 (parochial altruism)」が生じやすい理由は容易に説明できる²⁹⁾。この理論では集団同士が（意図的ではないにせよ）生存競争を繰り広げることが仮定されているので、内集団をひいきする性向および外集団に対して攻撃する性向は集団が生き残る確率を大きくし、ひいては個体の適応度を高めることになる。

27) Phelps et al. (2000) は、fMRI（機能的核磁気共鳴断層画像法）を使い、人種に基づく無意識レベルの偏見が扁桃体 (amygdala) の活性化と関連していることを示している。

28) 古典的な研究として、ヘンリ・タジフェル (Henri Tajfel) による「最小条件集団実験」が挙げられる (Tajfel 1981)。そこでは、社会的カテゴリー化を経るだけで、区別の根拠がなくても偏見が生じうることが示されている。

29) 偏狭な利他性に関する初期の文献として、Bernhard et al. (2006)、Choi and Bowles (2007) などを参照。最近のレビューとしては Rusch (2014) と Yamagishi and Mifune (2015) がある。

先にも触れたように、人間社会には他の生物にはない特殊な事情があるため、集団選択の余地が大きくなる。つまり、人間社会では文化を共有する集団が複数存立しており、集団内部で特定の行動様式が継承される。集団内の協力関係は制度によって強固に維持される一方で、集団間の軋轢が強い。そのため、集団レベルの淘汰圧が大きくなり、たとえ短いタイムスパンであっても生物の進化に匹敵するほどの進化プロセスが見られるかもしれない (Boyd and Richerson 1985; Richerson and Boyd 2005)。

さかんに引用される具体例として、文化人類学者レイモンド・ケリー (Raymond C. Kelly) が詳細に報告したヌエル (Nuer) 族とディンカ (Dinka) 族の事例が挙げられる (Kelly 1985; Henrich 2004)。両民族とも現在の南スーダンに居住しており、もともとはお互いに隣接した地域で類似の環境のもと牧畜生活を営んでいた。ところが、1820年以降の100年間でヌエル族は急速にテリトリーを拡大し、ディンカ族を圧倒する勢力をもつようになった。その理由として考えられるのは、それぞれの民族が採用していた政治的・経済的システムの違いであった。すなわち、ヌエル族はディンカ族と比べて組織化されており、多数のメンバーの生活を支えられるだけの生活形態を有していたのである。拮抗する集団が強い淘汰圧にさらされ、文化的・制度的な特性が一方の集団にとって有利に働いた例と言えよう。

果たして利他性は必然的に「偏狭」、言い換えると、内集団の人たちしか利他行動の対象となれないのだろうか。マルチレベル選択の理論を支持している人の多くはそのように考えているようである。というのも、集団選択が作用して利他性が進化するためには、グループが小規模でメンバーが移動する確率が十分に低いことや、誰がどの集団に属しているかが識別可能であることが必要であり、さらに、内集団びいきと外集団への攻撃性があれば (集団間の競争を生き残る確率が高まるため) 利他性が母集団の中で広がる公算は大きくなるからである。

しかし、この「偏狭な利他性」の議論を現実の社会現象の説明にそのまま使うことには慎重であるべきだろう。第一に、「偏狭な利他性」の概念を用いる論者は内集団びいきと外集団への攻撃性をセットで捉えており、これら2つの性向は相関しているとも主張している (Bowles and Gintis 2011: 140)。だが一方で、内集団の人に対して協力的な人は外集団の人に対しても協力的だということを示す実験結果もある (Yamagishi and Mifune 2009, 2015)。「偏狭さ」が出てきやすい状況とそうでない状況について、より詳細な検討が望まれよう。

第二に、「偏狭な利他性」が進化してきた環境の例として (特に狩猟採集社会であった時代の) 集団間の戦争が挙げられるが (Bowles 2004, 2009)、実際の戦争はそう単純ではなく、集団選択のモデルで描写されるような事態はあまり多くは起こりそうにない

(Rusch 2014). 競争を抑制して平和裡に共存するのがお互いにとって利益になる場合も多いだろう. また, 勝利によって得る利益やインセンティブは同じ集団内部であってもさまざまに異なりうるという点も考慮しなければならない. 要するに, 抽象的なモデルと具体的事例をつなぐための道具やブロックが足りていないのである.

2. 病理的利他性

「偏狭な利他性」は理論的には疑問なしとしないが, 利他性の危うさを端的に表している. ここで言う「利他性の危うさ」とは, 利他性はときに過剰になったり, 結局は他の人たちの利益にもならない結果を招来したりする, ということを指す.

具体的な人間の心理と行動を研究している人たちが関心をもつのは, 進化プロセスで獲得された遺伝子レベルでの「利他性」がどのような形で心理メカニズムに反映され, そして現在の人間にとっての(個人レベルまたは集団レベルでの)「利他性」とどのように関連しているのか, であろう. 換言すれば, (i) 生物的利他性と心理的利他性との間にいかなる連関があり, (ii) 心理的利他性がどのように作用して実際の行動や意思決定につながっていくのか, ということである. (i) と (ii) のどちらについてもすでに膨大な議論がなされているが³⁰⁾, ここでは (ii) に関するダークサイドについて述べておきたい.

心理のうえでは利他的, つまり他者のためになると思われることをしているが, そうした意図に反し, 予想を上回るマイナスの影響が自分または他者に及んでしまう, という場合がある. この現象は「病理的利他性 (pathological altruism)」と名づけられ, いろいろな分野の研究者がそのメカニズムや発生条件の解明に取り組んでいる (Oakley et al. 2011; Oakley 2013). もちろん, 「病理的」という言葉には価値判断が含まれている. 病理的であるか否かが判断される基準は絶対的なものではなく, 同じ現象であっても時と場合によって判断は変わりうる³¹⁾.

以下, 社会規範と関わる病理的利他性の例を2つ挙げよう. どちらの例も, 利他性の傾向が社会における相互作用を通じて増幅される可能性があることを示している.

30) 概観として, Batson (2011), Wilson (2015) を参照. 前者は利他行動を長年研究してきた社会心理学者の著書であり, 他者への共感が個人の利他行動に先行していることが実験によって確認されている. 日本語文献としては小田 (2011) を参照.

31) 病理的利他性の概念がどれほど有効であるのか, そもそも社会科学上の概念として適切なかどうかに関しては, 意見の一致があるわけではない. Oakley et al. (2011) に収められている諸論文を参照. 筆者自身は, pathological という形容詞は不必要にネガティブな印象を与え, しかも根底にあるアイデアを明確にしてくれるわけでもないため, 別の形容詞を用いるのが適切なのではないかと考えている (具体的に何が適切な形容詞かは, 指示したい現象が何なのかに依存することになるだろう).

(1) 行き過ぎた献身

互恵的利他主義とマルチレベル選択は、利他性の性質について異なる含意を導く。互恵的利他主義の場合は利他性が条件つき、つまりお返しをしてくれる相手にのみ利他行動が見られるようになる。それに対してマルチレベル選択の場合、内集団の人たちだけを相手にする限り、相手の行動についての期待によらない無条件の利他性が進化すると予測される。そこには、利他性の発動を制約するための原理は存在しない。

所属集団に向けた利他性ないし利他行動がエスカレートし、全体としては損失が利益を凌駕すると考えられる例は、枚挙に暇がない。戦時などの危機的状況でよく見られる自己犠牲、企業や学校に対する過剰な忠誠心、家族内部での共依存、所属集団の人たちに迷惑をかけたという自責の念からの自殺など、時代や状況によって姿を変えながら、行き過ぎた利他性によって深刻な弊害が生じる事例はたびたび発生する。

このような現象の背後には、パーソナリティ要因と環境要因の両方が作用していると考えられる。つまり、個人の共感性の高さに起因する場合もあれば、個人を取り巻く環境の変化（特に内集団を危機的状況に追いやるような変化）に起因する場合もある。

個人の利他性は、自ら利他行動をとるという現れ方をするだけではない。利他的処罰の例に見られるように、利他行動を他者に要求するという形で表面化することがある。そうになると、自発的に利他行動をとるつもりがなかった人でさえも、同様の行動を選択せざるをえなくなるかもしれない。

さらに進んで、そのような他者の要求が自己の内なる目として作用すると、やはり「自発的」に利他行動が選択されるだろう。多くの人間は他者の視線を考慮したうえで意思決定を行う。それゆえ、他者との相互作用によって、利他行動の頻度はしばしば増幅されることになる。

(2) 政策の利他性バイアス

(1) は個人レベルで利他性が過剰になる場合であったが、集合的な意思決定が利他性を過度に重視する方向になびく場合もある。特に、利他性を重視するあまり、解決すべき問題に適合しない法律が作られるということがある。少なくとも次の2種類のケースがありうる。

一つ目は、立法に携わる人々が特定の集団の人たちの利益だけを重く見るために（言い換えると、特定の集団内だけで利他性が働くために）、全体の利益にそぐわない法律ができる、というケースである。立法過程で利益集団が影響力をもつ国では、人々の厚生を不均等に反映した法律が成立しやすくなる。

二つ目は、利他性に基づく議論が先走った結果、実効性の乏しい立法や非効率な結果を

招く立法を制定してしまう、というケースである (Rubin 2014)。このケースについてはやや詳しく述べておこう。

歴史上顕著な例は、アメリカ合衆国で 1920 年に施行され、1933 年に廃止された禁酒法 (合衆国憲法修正 18 条、およびそれを受けて定められたヴォルステッド法) であろう³²⁾。禁酒法は突然できたわけではなく、1840 年代から脈々と続いてきた禁酒運動を背景としている。禁酒運動は主として敬虔なプロテスタントの人々によって牽引され、社会改善や貧困解消を目指す運動とも連動していた。その意味で、禁酒法に向けた動きも利他性、しかも相当に純粋な利他性に支えられていたと言える。

禁酒法に限らず、立法に対する支持の大きさは、当該立法がどのような帰結をもたらすかではなく、どのような意図のもとに作られているかに左右されがちである。立法、特に規制法を政治的に正当化する理由としてよく用いられるのは、①効率性 (法規制がなされれば人々の厚生が増進してより効率的な状態になる)、②利他性 (法規制がなされれば恵まれない人々が利益を享受する)、③危機感の緩和または解消 (法規制がなされれば損害の発生が回避される) である (Rubin 2014: 173)。だが、一般に人々は効率性の考え方になじんではおらず、法規制の帰結を評価することが不得手である。複雑化した社会では、効率性による評価はいつそう困難となる。

立法者の意図というのは、その法律に対して支持を与えるか否かを定めるうえで便利な基準として用いられやすい。ある人々 (架空であってもよい) に対する利他性が立法者の意図の中に含まれているように見せかけることができれば、人々の支持が得られる可能性は高まるだろう。

このように、政策決定の場面では利他性は力もちやすく、その結果、効率性を軽視した政策や非現実的な政策が採用される事態がときおり生ずる。

3. 小括

この節では、利他性が弊害となりうる場面を見てきた。現在の社会は多かれ少なかれ利他性によって動かされている面があり、したがって利他性は社会にとって欠くべからざる構成要素だと言える。しかし、内集団と外集団の境界がいったん確定すると、利他性は内

32) 合衆国憲法修正 18 条は、アルコール飲料を飲用目的で製造・販売・輸送することを禁止していた (アルコールの摂取は禁じられていなかった)。その後制定されたヴォルステッド法 (Volstead Act) は、0.5 パーセント以上のアルコールを含む飲み物を規制する内容であった。しかし、密輸に関連する犯罪の増加など、禁酒法に対する批判は次第に増加し、1932 年には禁酒法改正を訴えたフランクリン・ローズヴェルトが大統領選に勝利する。そして 1933 年、修正 18 条は修正 21 条により廃止されることになった (同時に、ヴォルステッド法は違憲となり効力を失った)。

集団の中で無条件性を帯びるため、逆機能をもつことがある。

利他性の逆機能については、「地獄への道は善意で敷きつめられている」という言葉（諺？）に見られるように、おそらく古来より感づかれていたであろう。ただし、利他性の逆機能の研究はその社会的重要性のわりにはまだあまり進んでおらず、体系的になってもいない。とはいえ、ジレンマ・モデルに固執していると見落とししやすい点を明らかにしており、社会規範の研究に対して有益な示唆を与えているように思われる。

IV. 社会規範における「利他性」の発現

1. 発現の諸形態

これまでに述べてきたところを踏まえつつ、社会規範と利他性の関係について若干の検討を行い、利他性の概念から導き出される別の視点を記しておくことにしたい。

(1) 望ましい要素としての「利他性」

ジレンマ・モデルに基づいた研究によると、社会規範の生成・維持の場面において利他性は大別して2つの局面で機能している。ひとつは、個人が社会規範に沿った行動をとるか否かという局面である。通常のジレンマ・モデルでは、社会規範が成立している状態とは協力行動が均衡として維持されている状態を指す。自己が利益を得る機会を放棄して他者の利益になるような行動をとる個人は、サンクションによる脅しとは無関係に協力行動を選択するだろう。社会規範を遵守している意識はないかもしれないが、結果として社会規範が指示する行動に従事することになる。

もうひとつは、社会規範に反した行動をとった他者を処罰するか否かという局面である³³⁾。実験から得られた知見によれば、誰かが協力行動をとらなかった場合、自分が直接の被害者であろうとなかろうと、その逸脱者に対する処罰行動が見られる。処罰行動を公共財と捉えるとすれば、これも自己利益を放棄して他者の利益になる行動だと解釈できる³⁴⁾。

(2) 望ましくない要素としての「利他性」

ジレンマ・モデルでは、利他性は望ましい性質（多くの場合、「効率的な状態をもたらす」

33) したがって、第一の局面に対してメタ・レベルに位置する局面である。

34) 処罰は単なるスパイト行動（意地悪行動）ではないのか、という考え方もある。スパイト行動とは、自分にも他者にも不利益になるような行動のことである。処罰が利他行動になるかスパイト行動になるかは、どのようなタイムスパンで見るとよくなるかによるだろう。

という性質)として描かれやすい。しかし前節で概観したとおり、利他性は不利益を増大させる逆機能を有することがあり、つねに望ましい性質だというわけではない。

マルチレベル選択の理論から予測されるのは、自分と同じ集団に属しているメンバーにとって利益となるような行動をとるというタイプの利他性の進化である。社会的カテゴリーによる他者の認識が人間にとって不可避なのだとすれば、内集団と外集団の区別は個人の意思決定に対して頻繁に影響を及ぼすようになるだろう。

内集団と外集団の区別が際立つ (salient になる) 状況では、先ほど述べた2つの局面——個人が社会規範に沿った行動をとるかどうかどうかという局面と、逸脱者を処罰するかどうかどうかという局面——における利他性は過剰なものになりやすい。つまり、社会規範を守りすぎたり、逸脱者を処罰しすぎたりする。しかもその利他性は歯止めがきかない場合があるため、全体として著しい非効率性を生み出すおそれがある。

閉鎖的な社会での処罰の激化³⁵⁾ やインターネット上の私刑 (サイバー・リンチ) など、社会規範のエンフォースが過剰になる事例は数多く報告されている。進化論の文脈ではこれは利他性の一形態にほかならないが³⁶⁾、協力規範の研究がこれらの事例を参照し、モデルで説明される利他性との関係を緻密に論じることはさほど多くない。

ジレンマ・モデルによる協力規範研究にもまだ解かねばならない問題はあるが、次項以下では、別の視点から社会規範と利他性の関係を素描する。すなわち、社会規範がどのように利用されるかに着目する視点である。もとより雑駁な素描ではあるが、もしかすると利他性の個人間ダイナミクスのありようを記述するのにいくらか役立つかもしれない。

2. 正当化のゲーム

通例、社会規範は「個人が選択しうる行動の範囲を制約するもの」として捉えられる。「～してはいけない」という禁止ルールであれ、「～するのが望ましい」という奨励ルールであれ、人々の行動を一定の領域内に制限するのが社会規範の特徴である。従来の研究では、社会規範のこの側面 (制約条件としての側面) に注意が払われることが多かった。人々の行動に焦点を当てて分析する限り、制約条件としての側面を切り出すのは致し方ないことではある。

しかし、社会規範にはこれとは違った側面がある。第一に、社会規範は他者を説得する

35) たとえば、炭鉱労働者の苛酷な生活を克明に描いた文献として山本 (2011 [1967]) を参照。

36) 自分自身の応報感情を満たすものであったとしても、処罰するという行為は適応度の観点からすればコストであることに変わりはない。つまり、「生物的な利他性」と「心理的な利他性」は別である。したがって、処罰が利他性をもつからと言って、主観的に「これは他の人のためだ」と思いながら処罰している必要はない。

ための材料として用いられる。たとえば、自分の利益とは相反する利益を求めている他者に対して、社会で通用している社会規範を援用しながら自らの正当性を主張するということがある。

第二に、社会規範は自己の行動を自分自身に対して正当化するために（換言すれば、自己イメージを傷つけないようにするために）使われることもある。組織ぐるみの不正行為（企業による虚偽報告や法令違反など）は、そのような正当化が如実に表れる例だと言えよう。「組織のために行動するのが望ましい」という規範があると、個人が不正行為に手を染めるのはたやすくなるかもしれない。解釈によっては不正行為を支持しうるような社会規範が存在することで、自己イメージを保ったままその行為を選択することが可能になるからである（Ariely 2012）。

道徳心理学の分野では、人々は道徳原理を一貫しない形で用いており、自分の欲求ないし選好を満足させるために道徳原理を「利用」している、という事実が複数の実験によって確認されている（Uhlmann et al. 2009; Haidt 2012）。人々は抽象的な道徳原理を自分の都合に合わせて解釈することがあり、その場で支配的な道徳はそのつど「創造」されているのである。

このことは、道徳に限らず社会規範にも当てはまるであろう。この点が視野に入ると、ただ制約条件とされることの多かった社会規範が、「制約条件としての性質」と「合理的な思考プロセスの産物としての性質」を併せもった複合的な現象として立ち現れてくる。

さて、社会規範が正当化のための道具だとすると、利他性の点から説明しやすい社会規範（誰かの利益になることがわかりやすい社会規範）が採用される可能性が高くなると考えられる。なぜなら、III 2(2) で触れたように、利他性を根拠とする主張は心理的・感情的な訴求力が大きく、対立する主張を封じられる確率が大きいからである。そしてまた、利他性を強調した社会規範を根拠にすれば、自己イメージを保ちながら自己の利益を正当化することも容易になる。

例として、「大金を使って贅沢品を買う」という行為を考えてみよう。この行為は時代によって評価が変わってきた。もし「消費欲の上昇から需要が増大し、それゆえ供給も増大することになるので、このような行為は経済発展を促し、他者の利益に結びつく」と説明できれば、贅沢をするという行為は一般の承認を受けやすくなるだろう。現代の先進国の社会では、経済政策にも表れているように、贅沢をするという行為は十分に正当化されていると言ってよい。

しかし、この行為が肯定的に捉えられるようになったのは、市場の働きが理解された後のことである。歴史上、贅沢をして享樂的な生活を送ることは悪い行為とみなされるほうがずっと多かった。たとえばキリスト教では、贅沢は七つの罪源のひとつである「傲慢」

に含まれるとされた。さらに、東洋でも西洋でも奢侈禁止令は何度も発布されている。

利他性の点から許容されるようになる行為とは逆に、非難の対象になったり禁止されたりする行為もある。たとえば、予防接種を受けないという選択は、他の人たちに与える影響を考えると非難の対象とされやすいであろう。実際のところ、予防接種にどのような効果および副作用があるかは理解が難しく、接種を受けるか否かについての決定の適切さは簡単には判断できないはずである。なかには接種を受ける理由が見出しがたいケースもある³⁷⁾。ところが、(予防接種を推奨する人たちの主張に見られるように) 予防接種を受けないことが利己的行動としてフレーミングされてしまうと、接種を受けないことを正当化するためには、より積極的な理由が要求されるかもしれない。

なお、利他性の「他」がどのような範囲の人たちを含むと想定されるかはここでも重要である。一般的に言って、「他」に含まれる人たちの範囲が広ければ広いほどよいというわけではない。問題になっている事柄によって「他」の範囲に差異はあるが、いずれにしても、関係するすべての人たちを包含することはむしろまれである。

3. 社会規範の「外面化」？

自分の主張を正当化するために社会規範を援用するとき、その社会規範は実態を反映したものであるとは限らない。要求水準が高すぎる等の理由で実際には人々に遵守されていない社会規範であっても、主張の正当化のために用いられる場合がある。この場合、一種の虚構(フィクション)が用いられているということになる。

虚構は虚構なりの生命をもつ。もし社会規範が自己利益に資するものであったり、複雑で無秩序な世の中をすっきりと理解する手助けをしてくれるものであったりすれば、たとえそれが実態を反映していないとしても、社会の中を生き延びることになるだろう³⁸⁾。このように用いられる虚構は決して脆くはない。

日本社会を少なからずコントロールしてきた「義理」の規範は、そのひとつの例と見ることができる。義理の意味内容については諸説あるが³⁹⁾、もとは好意に対する返しとして存在していた(源 2013 [1969]: 80)。義理は自発性を基礎としていわれていると言われており、初期の頃は「情的紐帯をもって結ばれ」ている人たちの間で成り立つ現象であった(源

37) 2016年1月現在定期接種とされているBCGの予防接種も、どの程度の効果があるのか、費用に見合った便益があるのか、他の方法と比べて効果の面でどのように異なるのか、わかりやすいとは言えない。ちなみに、厚生労働省は定期接種の中止も検討している(日本経済新聞2016年1月25日付の記事より)。

38) このプロセスを「ミーム(meme)」の概念や文化的進化のモデルを使って説明することも可能だろう。これらの概念についてはRicherson and Boyd(2005)、太田(2000: 25-57)を参照。

39) 義理について論じた代表的な文献として、本文中に挙げたもののほか、川島(1951)、安田(1974)などがある。

2013 [1969] ; 原 1983). 要するに自然な感情から湧き出るのが義理であったのだが、江戸時代の幕藩体制が整う時期以降になると、外的規範としての性格が色濃くなっていく。

井原西鶴や近松門左衛門らの文学作品には、義理が繰り返し取り上げられている。彼らの義理は現実の社会規範を描写したものではなく、「作者によって虚構化された作為的な状況をできるだけ違和感なしに享受者に受け入れさせるという上で、極めて大きな有効性を発揮し得る恰好な道具立ての一つとして採用されるに至った」(原 1983: 307)。やがてこうした作品は処世訓としての役割を担い、文学作品の中で示される義理の規範は農民の間にも広まっていった(桜井 1961)。

近世の封建社会における共同体は、凝集性の高い継続的關係によって特徴づけられる。そのような共同体では、通用すると人々が思っている社会規範は収斂しやすくなる⁴⁰⁾。正当化に用いられる社会規範の幅は狭くなり、他者の行動についての期待が確立していく。この段階になると、従うことを自己に対して強制する拘束力を各人が感じることになるであろう。

元来パーソナルな関係で自発的に生じていた「義理」が、外的拘束性のある「主君への義理」や「家制度にからまる義理」——つまり、公的支配原理と結びついた「義理」——へと変質することになった。源了圓はこれらを「私的世界の義理」および「公的世界の義理」という概念を用いて対比させており、近松以後の浄瑠璃に後者が明確に表れていることを指摘している(源 2013 [1969] : 182-195)。

心理メカニズムの点から言えば、義理の感覚は利他性の表出の一例である。進化プロセスによって形成された利他性が社会の相互作用を通じて外的拘束性を獲得し、個人からすると非自発性あるいは不本意性を伴う制約条件としての社会規範が現出するようになる。具体的な社会規範の内容は、特定の社会構造の中でいかなる正当化が有効であるかによって違ってくるはずである。

ここでは、規範の「外面化(externalization)」とでも呼ぶべき現象が生じている。つまり、本来は人々の自発性に任されていた行為が、正当化のゲームを重ねるうちに自発性を剥ぎ取られ、当然行うべき行為に転化していく。その結果として、自発的な利他性よりも程度の大きな利他行動が規範となり、これは別の正当化のしかたによって覆されるまでは社会的な力として作用し続ける。

社会規範の研究では、社会的に形成された規範をどのように個人が内面化するかが主題とされることが多い。この種の問題設定のもとでは、社会規範はすでに社会に存在しているものとされ、個人は受動的な立場に置かれるのが普通である。しかし、社会規範は人々

40) こうした特徴をもつ共同体は共有知識(Lewis 1969)が成立しやすく、社会規範は一定の幅に収まりやすい。

が紡ぎ出すものであり、内面化という一方向のみでは社会規範の動態は捕捉しきれないように思われる。

「利他性」を生む心理メカニズムが生得的に備わっており、かつ、利他性の感覚と合理性とが交錯して個人の意思決定や行動に影響を及ぼしているのだとすると、内面化の方向とは逆の「外面化」の議論を深化できる可能性がある。こうした議論は、Iで述べた合理性を前提としたモデルを補完し、現実の社会規範の態様をより厚く記述するための基礎を提供してくれるであろう。

V. おわりに

本稿は協力規範に関する研究のレビューから出発し、「利他性」の作用のしかたを概観したのち、次のように論を進めてきた。「利他性」が進化してきた過程を考えると、社会にとって望ましい現象だけでなく、望ましくない現象も容易にもたらしうる。「利他性」は社会規範の形成に必要な要素であるが、社会における人間の相互作用を通じて、従来の社会規範研究が想定していたものとは別の形で発現することがある。本稿ではこれのプロセスを「正当化のゲーム」として素描し、最後に規範の「外面化」について述べた。

ここで描くことができたのは大雑把な構図だけであり、正当化のゲームと規範の外面化が具体的にどのように展開するのかに関する詳細な検討は、いまだ進行中の課題である。また、「利他性」の概念をはじめとして、これらの展開過程を分析するための概念を整理し精緻化していくことも今後の課題となるだろう。

自然科学系の研究と社会・人文科学系の研究を取り結んで有意義な知見を生み出すためには、埋めなければならない溝がまだまだたくさん残されている。これらの溝をいくばくかでも埋めることができれば、現代社会の問題に対して多くの示唆を与えられるだろう、と私自身は考えている。本稿はそのような試みのひとつにすぎない。

社会をしばしば支配する重苦しきの根源が何なのか、私たちはなぜお互いに首を絞めあうことがあるのか、そして何より、どうすれば事態を打開してより良い方向へ進められるのか。こうした問題を探究するうえで、この小論や今後の研究がどなたかにとって少しでも参照に値するものとなればと願う。

参考文献

- Akerlof, G. A. and R. E. Kranton (2005) "Identity and the economics of organizations," *Journal of Economic Perspectives*, Vol. 19, pp. 9-32.
- Akerlof, G. A. and R. E. Kranton (2010) *Identity Economics*, Princeton University Press. [山形浩生＝守岡桜訳『アイデンティティ経済学』(東洋経済新報社, 2011年)]
- Alexander, R. D. (1987) *The Biology of Moral Systems*, Aldine.
- Anderson, C. M. and L. Putterman (2006) "Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism," *Games and Economic Behavior*, Vol. 54, pp.1-24.
- Ariely, D. (2012) *The (Honest) Truth about Dishonesty*, Harper Collins. [櫻井祐子訳『ずる』(早川書房, 2012年)]
- Axelrod, R. (1984) *The Evolution of Cooperation*, Basic Books. [松田裕之訳『つきあい方の科学』(ミネルヴァ書房, 1998年)]
- Basu, K. (1998) "Social Norms and the Law," In Newman, P. (ed.) *The New Palgrave Dictionary of Economics and the Law*, Vol.3, pp. 476-480, Macmillan Reference.
- Batson, C. D. (2011) *Altruism in Humans*, Oxford University Press. [菊池章夫＝二宮克美『利他性の人間学』(新曜社, 2012年)]
- Bernhard, H., U. Fischbacher, and E. Fehr (2006) "Parochial altruism in humans," *Nature* Vol. 442, pp. 912-915.
- Bicchieri, C. (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge University Press.
- Bourke, A. F. G. (2011) *Principles of Social Evolution*, Oxford University Press.
- Bowles, S. (2004) *Microeconomics: Behavior, Institutions, and Evolution*, Princeton University Press. [塩沢由典＝磯谷明德＝植村博泰訳『制度と進化のミクロ経済学』(NTT出版, 2013年)]
- Bowles, S. (2009) "Did welfare among ancestral hunter-gatherer groups affect the evolution of human social behaviors," *Science*, Vol. 324, pp. 1293-1298.
- Bowles, S. and H. Gintis (2011) *A Cooperative Species: Human Reciprocity and Its Evolution*, Princeton University Press.
- Boyd, R., H. Gintis, S. Bowles, P. J. Richerson (2003) "The evolution of altruistic punishment," *Proceedings of the National Academy of Sciences*, Vol. 100, pp. 3531-3535.
- Boyd, R. and P. J. Richerson (1985) *Culture and the Evolutionary Process*, University of Chicago Press.
- Brandt, H. and K. Sigmund (2004) "The logic of reprobation: assessment and action rules for indirect reciprocity," *Journal of Theoretical Biology*, Vol. 231, pp. 475-486.
- Brennan, G. and P. Pettit (2004) *The Economy of Esteem*, Oxford University Press.
- Choi, J.-K. and S. Bowles (2007) "The coevolution of parochial altruism and war," *Science*, Vol. 318, pp. 636-640.
- Dawkins, R. (1976) *The Selfish Gene*, Oxford University Press. [日高敏隆訳『利己的な遺伝子』(紀伊國屋書店, 1991年)]
- Ellickson, R. C. (1991) *Order without Law: How Neighbors Settle Disputes*, Harvard University Press.
- Fehr, E. and U. Fischbacher (2004) "Third-party punishment and social norms," *Evolution and Human Behavior*, Vol. 25, pp. 63-87.
- Fehr, E. and S. Gächter (2002) "Altruistic punishment in humans," *Nature*, Vol. 415, pp. 137-140.
- Haidt, J. (2012) *The Righteous Mind*, Pantheon. [高橋洋訳『社会はなぜ左と右にわかれるのか』(紀伊國屋書店, 2014年)]
- Hamilton, W. D. (1964) "The genetical evolution of social behavior, I & II," *Journal of Theoretical Biology*, Vol. 7, pp.1-52.
- Haslam (2006) "Dehumanization: an integrative review," *Personality and Social Psychology Review*, Vol. 10, pp. 252-264.

- Henrich, J. (2004) "Cultural group selection, coevolutionary processes and large-scale cooperation," *Journal of Economic Behavior and Organization*, Vol. 53, pp. 3-35.
- Kandori, M. (1992) "Social norms and community enforcement," *Review of Economic Studies*, Vol. 59, pp. 63-80.
- Kelly, R. C. (1985) *The Nuer Conquest: The Structure and Development of an Expansionist System*, University of Michigan Press.
- Lewis, D. (1969) *Convention: A Philosophical Study*, Harvard University Press.
- McAdams, R. H. (1997) "The origin, development, and regulation of norms," *Michigan Law Review*, Vol. 96, pp. 338-433.
- Neusner, J. and B. Chilton, eds. (2008) *The Golden Rule: The Ethics of Reciprocity in World Religions*, Continuum International Publishing Group.
- Nowak, M. A. (2006) "Five rules for the evolution of cooperation," *Science*, Vol. 314, pp. 1560-1563.
- Nowak, M. A. (2011) *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*, Free Press.
- Nowak, M. A. and S. Coakley, eds. (2013) *Evolution, Games, and God: The Principle of Cooperation*, Harvard University Press.
- Oakley, B. A. A. Knafo, G. Madhavan, D. S. Wilson, eds. (2011) *Pathological Altruism*, Oxford University Press.
- Oakley, B.A. (2013) "Concepts and implications of altruism bias and pathological altruism," *Proceedings of the National Academy of Sciences*, Vol. 110, pp. 10408-10415.
- Ohtsuki, H. and Y. Iwasa (2004) "How should we define goodness? Reputation dynamics in indirect reciprocity," *Journal of Theoretical Biology*, Vol. 231: 107-120.
- Ohtsuki, H. and Y. Iwasa (2006) "The leading eight: social norms that can maintain cooperation by indirect reciprocity," *Journal of Theoretical Biology*, Vol. 239, pp. 435-444.
- Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press.
- Ostrom, E. (2000) "Collective action and the evolution of social norms," *Journal of Economic Perspectives*, Vol. 14, pp.137-158.
- Phelps, E. A., K. J. O'Connor, W. A. Cunningham, E. S. Funayama, C. Gatenby, J. C. Gore, and M. R. Banaji (2000) "Performance on indirect measures of race evaluation predicts amygdala activity," *Journal of Cognitive Neuroscience*, Vol.12, pp.1-10.
- Posner, E. A. (2000) *Law and Social Norms*, Harvard University Press. [太田勝造監訳『法と社会規範』(木鐸社, 2002年)]
- Richerson, P. J. and R. Boyd (2005) *Not by Genes Alone: How Culture Transformed Human Evolution*, University of Chicago Press.
- Rubin, P. H. (2014) "Pathological altruism and pathological regulation," *Cato Journal*, Vol. 34, pp. 171-183.
- Rusch, H. (2014) "The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension," *Proceedings of the Royal Society B*, Vol. 281: 20141539.
- Sigmund, K. (2010) *The Calculus of Selfishness*, Princeton University Press.
- Sober, E. and D. S. Wilson (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Harvard University Press.
- Sugden, R. (1986) *The Economics of Rights, Co-operation and Welfare*, Basil Blackwell. [友野典男訳『慣習と秩序の経済学』(日本評論社, 2008年)]
- Tajfel, H. (1981) *Human Groups and Social Categories*, Cambridge University Press.
- Taylor, M. (1987) *The Possibility of Cooperation*, Cambridge University Press. [松原望訳『協力の可能性』(木鐸社, 1995年)]
- Tomasello, M. (2009) *Why We Cooperate*, MIT Press. [橋彌和秀訳『ヒトはなぜ協力するのか』勁草書房, 2013年]
- Trivers, R. L. (1971) "The evolution of reciprocal altruism," *Quarterly Review of Biology*, Vol. 46, pp. 35-57.
- Trivers, R. L. (1985) *Social Evolution*, Benjamin/Cummings. [中嶋康裕=原田泰志=福井康雄訳『生物の社会進

- 化』(産業図書, 1991年)]
- Uhlmann, E. L., D. A. Pizarro, D. Tannenbaum, and P. H. Ditto (2009) "The motivated use of moral principles," *Judgment and Decision Making*, Vol. 4, pp. 476-491.
- Warneken, F. and M. Tomasello (2006) "Altruistic helping in human and young chimpanzees," *Science*, Vol. 311, pp. 1301-1303.
- Warneken, F. and M. Tomasello (2009) "The roots of human altruism," *British Journal of Psychology*, Vol. 100, pp. 455-471.
- Waytz, A. and N. Epley (2012) "Social connection enables dehumanization," *Journal of Experimental Social Psychology*, Vol. 48, pp.70-76.
- West, M. D. (2005) *Law in Everyday Japan: Sex, Sumo, Suicide, and Statutes*, University of Chicago Press.
- Williams, G. C. (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton University Press.
- Wilson, D. S. (2015) *Does Altruism Exist? Culture, Genes, and the Welfare of Others*, Yale University Press.
- Yamagishi, T. and N. Mifune (2009) "Social exchange and solidarity: in-group love or out-group hate?" *Evolution and Human Behavior*, Vol. 30, pp. 229-237.
- Yamagishi, T. and N. Mifune (2015) "Parochial altruism: does it explain modern human group psychology?" *Current Opinion in Psychology*, Vol. 7, pp. 39-43.
- Young, H. P. (2008) "Social norms," In Durlauf, S. N. and L. E. Blume (eds.) *The Palgrave Dictionary of Economics*, 2nd ed., Palgrave Macmillan.
- 飯田高 (2004) 『<法と経済学>の社会規範論』勁草書房.
- 巖佐庸 (2015) 「協力の進化：人間社会の制度を進化生物学からみて」, 亀田達也編著『「社会の決まり」はどのように決まるか』(勁草書房), 15-47頁.
- 太田勝造 (2000) 『法律』東京大学出版会.
- 岡田章 (2011) 『ゲーム理論〔新版〕』有斐閣.
- 小田亮 (2011) 『利他学』新潮選書.
- 川島武宜 (1951) 「義理」『思想』327号, 21-28頁.
- 神取道宏 (2015) 『人はなぜ協調するのか』三菱経済研究所.
- 桜井庄太郎 (1961) 『恩と義理』アサヒ社.
- 中丸麻由子 (2011) 『進化するシステム』ミネルヴァ書房.
- 原道生 (1983) 「虚構としての『義理』」相良亨=尾藤正英=秋山虔編『講座日本思想 第3巻』(東京大学出版会), 289-328頁.
- 真島理恵 (2015) 「間接互惠性状況での人間行動」, 亀田達也編著『「社会の決まり」はどのように決まるか』(勁草書房), 117-147頁.
- 松井彰彦 (2002) 『慣習と規範の経済学』東洋経済新報社.
- 源了圓 (2013 [1969]) 『義理と人情』中公文庫.
- 安田三郎 (1974) 「義理について：日本社会論ノート」現代社会学1月号 179-197頁, 2月号 163-174頁.
- 山本作兵衛 (2011 [1967]) 『炭鉱に生きる』講談社.