

氏名	水野景子
学位の専攻分野の名称	博士（社会学）
学位記番号	甲社第77号（文部科学省への報告番号甲第815号）
学位授与の要件	学位規則第4条第1項該当
学位授与年月日	2024年3月16日
学位論文題目	社会的ジレンマ状況におけるサンクションの協力阻害効果の検討
論文審査委員	（主査）教授 清水裕士 （副査）教授 稲増一憲 教授 石田淳 犬飼佳吾（明治学院大学准教授）

論文内容の要旨

本論文は、社会的ジレンマ状況においてサンクション（行動に対して賞や罰を与えること）制度が協力的行動を阻害する可能性、およびそのときの意思決定プロセスを検討した論文である。本論文は8章によって構成されており、大きく分けて2つの研究が行われている。1つは、社会的ジレンマ状況においてサンクション制度を導入することで協力的行動が阻害されるのかを実証的に検討するための研究である。もう1つは、数理統計モデリングを用いて、サンクション制度における人々の意思決定プロセスを明らかにするための研究である。本論文の概要は以下の通りである。

序章

個人が各々の利益を追求した結果、集団として望まない帰結を生むことがある。このような個人の合理性が集団の非合理性につながる状況を表す理論的ツールに社会的ジレンマ（social dilemma）がある。社会的ジレンマは、(1) 個人に協力と非協力の選択肢があるとき、(2) 一人ひとりの個人にとっては協力を選択したほうが得であるが、(3) 全員が非協力を選んだ場合に、全員が協力を選んだ場合よりも集団全体の利益が小さくなっていく状況と定義されている。社会的ジレンマは身近な問題から大きな社会問題まで、多くの問題に共通する構造である。

社会的ジレンマ研究において、人々は、最初は協力し合うがしだいに協力をしなくなり、全体的に協力しなくなっていくことが知られている。このような問題に対して、サンクション制度の存在が有効であることが示されてきた。望ましい行動に報酬を、望ましくない行動に罰を与えることは、人の行動を変えるための手段としてしばしば用いられる。

しかし、サンクション制度の導入は、むしろ協力を妨げる可能性を内包していることが先行研究によって指摘されている。具体的には、長期的な協力の維持を考えた場合、他者への信頼の破壊や、他者に対する利他的動機の阻害などの弊害により、罰制度を廃止した後、導入しなかった場合よりも協力の程度が低下する可能性があると考えられる。本研究ではこれをサンクションの協力阻害効果と呼ぶ。もし実際にサンクションの協力阻害効果が起こる場合、サンクション制度の導入には慎重であるべきであり、このような状況を防ぐ制度設計の模索が不可欠である。したがって、サンクションの協力阻害効果についての検討が、協力をどのように維持するかという問題において取り組まれるべき課題であることは疑いようがない。

しかし、先行研究において社会的ジレンマ状況の協力阻害効果そのものを検出した研究は非常に少なく、この効果があるという前提のもとでさまざまな心理学的解釈が行われているという限界点もあった。そこで本研究の前半部分である2章から4章では、実験研究によってサンクションの協力阻害効果が実際に生じるのかを確認した。本論文ではサンクションはすべて非協力に対する罰として定義されているため、以下、罰の協力阻害効果と呼ぶ。

2章では、罰のない状況から途中で罰制度を導入し、その後廃止した場合に罰の協力阻害効果が起こるのかを検討した。実験では、社会的ジレンマの利得構造を持つゲームを参加者に行ってもらい、そこで参加者の行動を記録し、分析を行った。2章では罰がない状況を5回、罰制度を導入したゲームを5回、そして罰制度をとりはらったゲームを5回、合計15回のゲームを行った。参加者は実際に自身の選択によってもらえる報酬が変動するため、実際にインセンティブがある状況での行動が測定された。協力の阻害効果は、罰制度を導入する実験群と、罰制度を導入しない統制群の、罰制度がとりはらわれた11回目以降の5回のゲームの平均協力率の差として定義された。これは、罰制度がなくとも、ゲーム回数を重ねると協力が低下する現象があるため、罰制度前後の比較では正確な協力の阻害効果が検出できないためである。実験は、サンプルサイズと実験後に行うすべての分析プランをあらかじめ決定され、事前登録（事前に修正ができない状態でOSFに登録すること）がなされたうえで実験が行われた。分析の結果、罰制度による協力阻害効果は確認されなかった。

3章では、2章で得られた結果をもとに、いくつかの観点から実験設定を変更し、罰による協力の阻害効果を検討した。まず、2章で行われた研究は、協力阻害効果が生じたとされた先行研究と全く同じ設定ではなかった。具体的には、先行研究では実験条件では最初から罰制度が存在しており、途中から導入した2章の研究と異なっていた。そこで、先行研究の直接追試実験を十分な検出力によるサンプルサイズを決定し、事前登録を行った。その結果、やはり罰による協力の阻害効果はまったくみられなかった。続いて、社会的ジレンマ状況における罰制度は、先行研究では大きく分けて2つの観点で整理された。1つは罰の大きさ、もう1つは罰の種類である。罰の大きさは非協力によって得られるリターンを上回る罰の大きさであるか、上回らない程度の大きさであるかによってナッシュ均衡が変化する。また、罰の種類は最も協力率が低かった人を罰する競走罰と、一定の協力の程度を下回った場合に罰する規範罰の2種類が存在した。そこでこの2種類の罰について4つの組み合わせすべてにおいて、事前登録を行い、実験を行った。その結果、すべての状況について罰の協力阻害効果は認められなかった。

4章では、実験で用いられた社会的ジレンマゲーム状況がそもそも現実において問題とされるような、「協力が規範的に要請されている状況」ではないことが議論された。そこで、「協力することはよいことである」という教示を行った場合における罰の協力阻害効果を検討した。また罰制度の教示についても検討された。4章では罰制度と同じ利得構造であるが、非協力を選択するために自身がオプション料金を支払うという条件を設定した。しかし、4章においても罰の協力阻害効果は検出されなかった。

続いて、後半部分である5章から7章では、罰制度中の意思決定プロセスを数理モデルによって表現する研究が行われた。4章までの実証研究から、社会的ジレンマ状況における罰の協力阻害効果は、非常に小さいか、あるいは見られないことが明らかとなった。しかし、罰制度が導入されることで人々の意思決定プロセスは変化している可能性があるため、数理モデルによる検討が行われた。

5章では、他者に対する協力傾向のモデルである、社会的価値志向性を数理的に表現したモデルを用いて、社会的ジレンマ状況における意思決定モデルを導出した。具体的には、他者に対する利他性および不平等を回避する傾向を含めた武藤（2006）のモデルを前提に、社会的ジレンマ状況における他者がどれほど協力するかの期待をパラメータとして設定し、そこから協力と非協力それぞれの期待効用を導出した。そこから、人々は利他性が高い場合に協力すること、不平等回避傾向が高いほど、他者が協力するかどうかによって協

力するかどうかが変化することが導かれた。

6章では、利他性と平等性のパラメータを推定するためのあらたな課題を提案し、その方法が先行研究で使われている社会的価値志向性の尺度よりも優れていることを数理的、あるいはシミュレーション研究、そして実証研究によって検討した。これまでは社会的価値志向性はSVOスライダーとよばれる尺度で測定されていたが、測定上の問題点も先行研究によって指摘されていた。そこで、より利他性と平等性を測定するのに適した課題と、推定方法を新たに提案した。まずシミュレーション研究から、提案方法が理論的に仮定された値を正しく推定できていることを確認した。また2つの実験研究から、社会的ジレンマ状況における協力選択率を、SVOスライダーよりもよりよく予測できることを明らかにした。

7章では、5章で提案されたモデルと、6章で提案された測定方法を用いて、罰制度が存在する社会的ジレンマ状況の意思決定プロセスを、2章のデータを用いて検討した。その際、先行研究で想定されていた罰制度による信頼の破壊と、利他的動機の損失という2つのメカニズムに加えて、信頼や利他的動機は何も変化せず、罰制度導入前にリセットされるという信頼リセットモデルを検討した。ベイズ統計モデリングによるモデル比較の結果、信頼リセットモデルが最も予測力が高くなった。このことから、罰制度を導入することによって、先行研究で言及されてきたような、他者の協力期待の減少や、本人の利他的動機の減少などによっては行動を予測できず、どちらも変化しないモデルが最も予測できることを明らかにした。

8章の総合考察では、各章での研究結果をふまえて、本論文の結論および今後の研究への展望が論じられた。本研究では、社会的ゲームにおいて罰制度が一度導入された後に取り除かれた際の協力の程度について、すべての条件において、罰制度が取り除かれた後の協力が、罰制度のある状況を経験していない統制群と同程度になることを示した。当初の仮説は支持されなかったが、すべての研究を事前登録したうえで、合計1544名の大規模なデータを用いた包括的な分析によって、罰の協力阻害効果だけでなく協力を促進する効果も見出されなかったこそが、本研究の最たる貢献であり、新たな知見であることが主張された。

一方で、本研究はすべて実験研究であり、その一般化可能性については慎重であるべきであることが論じられている。現実社会は、他者の行動をすべて観察できるわけでもなく、また行動の選択肢が明示的に提示されるわけでもない。しかし、一方で本研究の知見は社会的ジレンマ状況における人々の振る舞いを理解するためのベンチマークとして機能しうる。そして統制された状況と現実場面における行動の違いがあったとき、その違いを知ることはさらなる人間行動の理解につながるだろうことが論じられた。

今回の研究が仮に現実社会における社会的ジレンマ状況に適用できた場合、政策立案や企業、教育機関などの組織内のルール設計において、罰制度は存在する間のみ一時的な効果を持つ可能性を考慮すべきであることが示唆される。感染症対策への非協力に対する罰金や、基準を超えた電力消費に対する追加料金の導入など、感染症が流行している期間や電力が逼迫している期間という特定の期間だけ協力的な行動を促進したい場合には有効であるが、一時的な罰制度による長期的な「協力の癖付け」を狙うことは効果的ではないことなど、本研究の知見についての社会制度に対する応用の問題について議論された。

本研究のアプローチは、人の意思決定や心の仕組みに関する再現可能な基礎研究の知見を積み重ね、社会に応用できる枠組みで人間理解を深めることを目指している。本研究のアプローチは、制度設計に関する他の社会科学分野の研究の基礎研究としての社会心理学研究の性質を十分に満たしていることが主張された。

論文審査結果の要旨

水野氏が提出した博士学位申請論文は、社会的ジレンマ状況におけるサンクションの協力阻害効果を実証的、数理的に検討した研究である。これまで先行研究で存在するとされていた協力阻害効果について、緻密な事前計画と大規模な実験データを用いて、少なくとも社会的ジレンマ状況においては生じないことを示し

た。また数理モデルによる意思決定プロセスの検討によって、先行研究で心理的に生じるとされていたモデルがすべて実験データを予測できないことも明らかにした。このように、大規模な実験による正確な因果推論と、そのメカニズムについての緻密な数理的な検討によって、社会的ジレンマ状況における罰制度の協力阻害効果の理解に大きく貢献した研究であると言える。

序論部分では、膨大な社会的ジレンマについての先行研究を、緻密にかつ本論文の射程としている領域について網羅的にレビューされている。その中で、これまでの研究がそれぞれ独自の観点で主張してきた協力阻害の心理現象を、制度導入における協力行動の変化と、その背景にある心理プロセスを分けて考えるべきであること、そして、まず前者の効果の確認が必要であることが的確に論じられている。本論文は非常によく構造化され、緻密に組み立てられており、論文構成についても高く評価することができる。

また、本研究は心理学分野だけでなく、他の周辺分野に対しても大きな貢献がある。上記のように協力の阻害効果を行動の変化と心理プロセスを分けて定義されていることで、実験研究部分は心理学だけでなく、社会制度に関心がある行動科学全般に対して大きなインパクトを持つ。また、意思決定メカニズムを行動経済学でも利用されている社会的価値志向性に基づいて数学的に導出していることから、経済学だけでなく、数理社会学にも波及する研究となっている。

このように、本論文は理論と実証の両側面においてすぐれた貢献のある研究であると言えるが、審査にあたった委員からは以下のような課題も指摘された。

1. 実験で用いられた罰制度の導入と、実際の現実場面での制度導入では、タイムスパンの違いなど、様々に異なりうることから、実験研究だけからそのまま制度に対する知見に結びつけるには一定の制限がある。
2. 数理モデルによって最も予測力が高かった信頼のリセット効果について、更に社会心理学的な解釈を掘り下げることができたのではないか。今回は実験の短いタイムスパンにおいてはこのモデルのようなメカニズムは成立しうるが、現実の制度では成り立ちうるのかについて、さらなる考察があるべき
3. 社会的ジレンマ状況における協力という狭い行動範囲をこえて、罰制度による自発性の喪失といった観点で、状況を拡張していくことでさらなるヒントが得られる可能性もある。

などが議論された。これらの指摘に対しても、水野氏は誠実かつ的確に回答していた。また、これらの指摘は本論文の論理的な破綻や実証研究上の問題点というよりは、今後の研究に期待される助言的な課題であると理解することが妥当である。

口頭試問の結果を含む、以上の審査結果にもとづき、本委員会は、水野景子氏が博士（社会学）の学位を授与されるに十分な資格を有していると全員一致で判断した。