

「囚人のディレンマ」考

——社会科学的思考訓練におけるモデルの効用——

三 星 正 臣

本稿は社会科学的思考の訓練に関する一つの試行の報告であり、かつまた試行の結果得られたデータに基づき、国際政治の力学や政治意識に関し、若干の学際的観察を行なうものである。

社会科学全般に共通する基本的な理念や手法を、初学者に教育するに当たっての、ある方法論上の提案を意図したものであるが、副次的に、国際政治論や平和論の意識調査方法についての示唆にも成り得ると考える。両者の視点を並行しつつ論を進めるのは論証上の破綻や混乱を招く恐れはあり、本稿でもこうした欠陥は十分以上に露呈しているのであるが、生成途上の学際的論究に伴う模索の一過程として許容願いたい。

1. 状 況

ゲーム理論で使われる「囚人のディレンマ」なる実験手法がある。設定は周知のものである。

重大犯罪の嫌疑で取り調べを受けている二人の囚人（正確には容疑者）がいる。共犯関係にあると検察は考えている。しかしながら犯罪の物的証拠は一切無い。自白を引き出すため、検察は一策を案じた。引き離されて収容されている囚人を、一人ずつ個別に呼出し、検察は取引を申し出る。取引内容は以下である。

- (1) 一日でも早く自白した囚人を釈放のうえ、多額の賞金を与える。
- (2) 自白しなかった（あるいは自白が遅れた）他の一人の囚人は死刑に処する。
- (3) 二人が同日に自白した場合、二人ともに懲役10年。
- (4) 二人が最後まで自白しなければ、二人ともに証拠不十分で釈放。

しかし検察は拘留期間を明らかにせず、最後に囚人の心理を攪乱させる一言を加える。共犯者が今日自白しないという保証があるか、よく考慮したうえ本日中に態度を決定するようにと、示唆するのである。

囚人は隔離されており、互いに相手の行動を予測できない。そこで自白すべきかどうか悩むことになり、これがディレンマというわけである。この際、囚人が無実か否かは関係ない。また自白のみでは有罪となし得ないという、近代法の原則も関係ないものとする。

設定はノン・ゼロサム関係であり、相手側とのコミュニケーションの欠如が加わる。上記の例

では、与件は極限状態に近く設定してある⁽¹⁾。対立関係にある複数の権力間の意志決定メカニズムを解説するにあたり、モデルとして簡明であり、初学者の興味も引きやすい。国際政治、国際経済、コミュニケーション論で用いられることがあるが、本稿では社会科学全般にわたり、初学者が社会科学の本質を認識する手段としての、このモデルの効用を論じたいのである。

設問の解は「自白する」である。相手の行動が未知数であり、個体の生存本能が優先するとの前提では、(A)釈放 or 死刑、(B)釈放+賞金 or 懲役10年 の選択肢のうち(B)を取るの論理的帰結（自然解）である⁽²⁾。論理上はもともと迷う余地はなくディレンマでもない。したがって双方が自白し、結果的に二人の囚人は懲役となる。初学者はここで先ず、社会科学が対象とする人間社会を、当為としてでなく現実として理解すべきことを学ぶ。

「囚人のディレンマ」の解はしかし一つではない。上記の論理的解に至るまえに、初学者が平常の感覚で、いわば「人間的に」望むのは殆どの場合、双方ともに自白せず釈放に至る選択であろう。それを「種」としての選択と呼ぼうが、「文明」として或は「人類」としての選択と考えようが自由であるが、この希望的感情的選択は、ひとたび視点を変えて囚人二人を価値の単位ととらえた時には第二の論理的解（協力学）にもなるのである⁽³⁾。人間個々にとっては論理的選択であるものが、個体を越えた社会にとっての論理的選択と矛盾する状況が提示される。「ディレンマ」はここに存在するのである。

ところで初学者は既に第一の論理的解（自然解）を得る過程で、社会現象を希望的観測によりいわば我田引水的に解釈する態度を修正されている。その結果彼らが直面した社会像は、力の論理によって動く *homo homini lupus* の世界である。永遠に繰り返される懲役を運命づけられた人類の姿である。その彼らに第二の解（協力学）を提示することは多少の混乱を伴う。ともすれば現実と、人間の尊厳を対比させる類の、精神論的な理解に逆行しがちである。

第二の解への回帰は論理的帰結として提示されなければならない。望むらくはこの論理的回帰は学習者自身によって発見されねばならない。このことで初学者は、社会科学が対象たる人間社会を *homo economicus* 或は *homo politicus* の集団として、冷徹に分析しなければならないこととともに、さらに進んで、社会科学の目的は、単なる分析ではなく社会の価値実現であり、実現に至る手段の探究であることを学ぶ。ちなみに、同時に複数の論理的解が存在することの理解は、pluralism 的思考への案内にもなる。しかし何よりも大切なのは、社会科学を、現実に立脚しつつ、しかし現実の分析に留まらず、さらに当為を目指す人間社会の学としてとらえること、即ち人間の望む結果を招来すべき論理的手段を探るのが、社会科学の目的であることを、初学者に認識させることである。

「囚人のディレンマ」の場合、実験を繰り返すことにより、第二の解を論理的解として認識させることができる。囚人の選択は、双方ともに自白しない場合を除き、一回かぎりの選択である

はずだが、実験では繰り返しを許容する。検察は前日の結果を囚人双方に告げ、あらためて今日の選択を迫る。繰り返すことにより、双方の選択にどのような変動が生じるかを探るのである。300回の実験の結果、判明しているデータは以下である⁽⁴⁾。

- (1) 一回目のテストで正解（第一の解）を出す被験者の率は、50%程度である。
- (2) テストを繰り返すことで正解率は高くなる。30～40回では73%に達する。
- (3) さらに繰り返すと逆に正解率は下がる。最後は23%に落ちる。

(1)から(2)の数値変化は第一の解に至る学習効果と見て良い。相手の裏切りにより、被験者は、人間社会の力の論理の実相を垣間見るのである。

(2)から(3)の数値変化が、第二の解の認識過程である。被験者は、実験を繰り返しかう着状態が続く間に、個々にとっての最善の選択が、結果として永遠の幽閉状態に過ぎないことを認識し始める。また偶然の間違った選択が、理想的結果をもたらす若干の例も経験する。この段階で変化への試行が始まる。相互の試行は相互に認識され、やがて試行から必然へと視点の変化をもたらす。個を越え、自己と他を含めた価値単位の再発見である。(3)はしたがって、第一の解の正解率下降ではなく、第二の解の正解率上昇としてとらえなければならない。

しかしながら300回の繰り返しテストは、演習ならともかく、初学者に対する講義においては実行不可能である。実際例でも、講壇上で数人の学生に数回の実験を行ない、以後の経過については既成のデータに基づく説明で代行させることが多い。したがって学生は、100回以上の繰り返しを要する(2)から(3)の過程は勿論、(1)から(2)の課程も十分に体験できない。第二の解の論理的帰結を、自ら発見する機会はないのである⁽⁵⁾。このため「囚人のディレンマ」第一の解発見の体験により、惹起された問題意識と知的関心を持続させつつ、さらに第二の解の発見を通じて、学生自身がいわば啓示のように、社会科学を当為を目指す学として、再びとらえ直す道は失われる。講壇からの解説だけでは、学生は理解するであろうが、説得力を持つとは限らない。社会科学との出会いにあたって、学の本質についての問題意識が、刷り込み効果のように強固に形成されることの重要性は言うを待たないであろう。それは研究者に限らない。一般教養課程で初めて社会科学に接する学生は、殆どが数年後に実務者としての長い人生に入る。彼らをいかに科学的に、主体的に、目的意識をもって、社会事象に正対する人間に育成するか、それは社会科学に携わる教育者の責務であろう。

2. 試 行 I

論者にとっての当面の課題は、かくして「囚人のディレンマ」第一の解、第二の解を迫る疑似体験を、実施可能な時間へ短縮する手段であった。試みに小グループによる討論を併用した。

被験者は18～19才の女子学生87人、論者の比較文化論の受講生である。論者は「囚人のディレ

ンマ」を、主として pluralism 的思考訓練の一過程として援用したのであるが、その論及は本稿では省く。被験者は社会科学専攻志望ではない。むしろ社会科学に対する関心は薄い。国際問題には若干の関心はあるが、理論面に対してではない。要するに社会事象の理論的分析は未経験であり、初学者の要件は満たしている。また「囚人のディレンマ」に対する予備知識も皆無である。一回目、通常の方法で行なわれた回答結果は下記の数値を示した。

自白スル 32 自白シナイ 55

囚人の状況の説明の後、数分間の思考時間を与え、無記名による回答である。正解率は37%であり、通常の50%より若干低いが、性別・年齢層などの偏りを考慮に入れば、許容範囲内としてよいであろう。

二回目、回答前に小グループによる討論を行なわせた。グループは語学のクラス単位とし、7人から21人までの人数不揃いの8グループ（グループA～Hとする）が成立した。各グループに与えた設定は次の通りである。

- (1) クラス全員（40人）がそれぞれ囚人の立場にあるものとする。それぞれの共犯者は、対抗クラスの学生であり、ひとりひとり対応して組み合わせられている。
- (2) グループのメンバーはクラスを代表して、回答を求められている。彼らの回答の比率が、クラス全員の回答比率となる。
- (3) 回答はグループ内で十分討論したうえ、グループの全員が個別に行なう。
- (4) 討論の結果メンバーの回答を拘束するか、自由回答とするかはグループの自由である。
- (5) 他のグループとは情報交換などの接触は行なわない。
- (6) クラス全員が他の一クラス全員と、無作為に組み合わせられ対抗戦を行なう。回答結果により、釈放された人数（生き残り率）が多いクラスを勝者とする。両クラス同人数の場合、両者とも勝ち上り、両クラス全員が懲役の場合、両クラスとも不戦敗。クラス対抗トーナメント戦で、勝ち残ることが最終目標である。最終勝者の数は、0から8クラスまでの可能性がある。

回答結果は下記である。回答にはグループ（クラス）名のみを記入させた。

	A	B	C	D	E	F	G	H	計
自白スル	5	4	8	15	15	6	5	9	67
自白シナイ	2	4	2	6	0	2	4	0	20

正解率は77%である。通常のテストを繰り返して到達する最高値にきわめて近い。グループE

及びグループHについては、自由意志による回答が許されなかったと想定されるので、その数値を除外してもなお、正解率は68%である。討論の参加を通じ、被験者は「四人のディレンマ」の第一の解の論理性を習得したと推定される。それは前述のとおり、彼らが現実を客観的に認識する科学的手法の第一歩を踏み出したことを意味する。

単に討論による情報交換と相互説得のみが、この過程を速めているのではあるまい。クラスの代表意識もまた、被験者の状況認識を鮮明にさせていると見られる。決定結果についての、他者への責任意識である。これも社会科学の初学者が、早期に習得すべき重要な要素であると、論者は考える。責任意識とは、自己を越えて広げられた価値単位の認識にほかならない。したがってそれは、第二の解に至る転機を内在している。自己から討論グループへ、さらにその場に存在しないクラス全員へと広がった単位意識は、やがて相手グループを含む意識に変化する可能性がある。必然性ではない。しかしこれが囚人が二人共に解放される唯一の可能性である。人類が幽閉状態から脱却する唯一の可能性なのである。

被験者の第一の解発見の醸成が、討論により加速されたことは疑いない。しかし第二の解の発見に関しては、上記テストの数値からは推測し難い。おそらく設定を変えて、再度の討論を経ることで、相当数の被験者に対して、300回繰り返しテストに近い効果を得られるであろうが、論者はまだその設定を案出していない。ここでは、上記のテストでも、既に一部に第二の解への移行が始まっていることを指摘するに留める。

グループBとグループGに注目したい。第一の解に関するかぎり、他グループに比し正解率は低い。論者は当初、討論による客観的認識が他グループに遅れているものと解した。全グループについて討論内容の聴取を行なった結果、この両グループでは、実は第一の解の認識を越えて、第二の解への模索が始まっていることが判明した。

グループB・Gの内部では、意見は完全に両分された。第一の解を固執する意見は、「他グループが第一の解を選択するのは絶対」とするものであった。したがって自滅を避けるため、自グループも第一の解を固執せざるをえない。これにたいし反対意見は、「第一の解を全グループが選択すれば勝者はいない。複数のグループが選択した場合も最終戦で勝者はいない。とすれば、必然的な不戦敗を選ぶより、他グループでも同様な討論が行なわれている可能性に期待したい」という主張である。意見は分かれ、分かれたままの自由回答となった。

反対意見の立論には誤りがある。複数のグループが第一の解を選択した場合、組み合わせの如何により、その中の一グループが最終勝者になる可能性はある⁽⁶⁾。第一の解の選択は「必然的な不戦敗」ではない。しかし重要なのはこのような戦術論ではない。

反対意見の主張には、全グループによる第一の解では、最終勝者はいないという認識がある。他グループが異なる選択を試みる可能性の認識がある。その二つの認識を、他グループと共有し

ている認識がある。

ここに第二の解に向かう流れの水源を見ることは容易であろう。

3. 分 析

被験者は、ゲームの意味についてオリエンテーションを与えられていないにも拘らず、既にして討論の中で「囚人のディレンマ」と国際政治のアナロジーを予感している。予感に即してゲームの結果を現実社会の事象に対比することは、被験者にとり分かりやすく、被験者の発見をより強固に定着させるであろう。被験者の「囚人のディレンマ」第一の解、第二の解の発見遍歴が、直ちに社会科学の方法論の理解と結びつくと考えるのは、初学者には過重な期待である。具体的社会事象を例証として媒介させるのが有効であろう。以下の分析は、この定着作用を期待して、前述のテスト結果を、国際政治等の局面にパラフレーズした例である。

87人の被験者は一回目のテストと二回目のテストで解答比率を逆転させた。

	テスト 1	テスト 2
自 白 ス ル	32人 (37%)	67人 (77%)
自 白 シ ナ イ	55人 (63%)	20人 (23%)

被験者には予め相手を特定していない。87人を任意に二人ずつ、囚人同志として組合せる。結果の比率は、理論上は下記のようになる⁽⁷⁾。

	囚人両者釈放の比率	両者懲役の比率	一方死刑一方釈放の比率
テスト 1	39.7%	13.2%	47.1%
テスト 2	5.1%	59.1%	35.8%

したがって、87人の運命は次のように分かれる。

	釈放	懲役	死刑	賞金付釈放	(釈放者計)
テスト 1	35人	12人	20人	20人	87人中55人
テスト 2	4人	51人	16人	16人	87人中20人

比較的信頼協力関係にある社会と、不信と非協力が利益対立を際立たせている社会に、テスト 1 とテスト 2 の状況を対比させることは理解し易いであろう。むしろ単純に過ぎて、数値的解

析の必要性もないかに見える。偶然ではあるが、「自白シナイ」の回答数と釈放者数も一致している。「自白シナイ」の回答数と釈放者数の増減は、一見算術級数的に比例しているようだが、この数値はしかし、釈放者同志が再度「共犯者」関係になった時、その実相を明らかにする。テスト1とテスト2で釈放された55人と20人が、それぞれ再び囚人として組み合わされるとする。全員が以前の選択を繰り返すと考えられる。その結果はテスト1の55人については、33人が再度釈放される。テスト2の20人のうち再度の釈放は3人ないし4人にすぎない。三回目には0となることは明白である。結果の差は、もはや算術級数的とは言えない。

囚人を国家に置き換え、対立抗争の国際社会を類推することは既に被験者の胸中にある。絶滅する国家、隷従する国家、生き残る国家を囚人の運命に当てはめたとき、人類が何度の大戦に耐えられるか、彼らが想像力を働かせるのも難くはない。

しかし「囚人のディレンマ」は、あくまでもゲームの理論である。それは多様な現実の一面を強調したモデルにすぎず、現実との照応には限界があり、想像や示唆の段階に留めておくべき仮説も多い。上記の分析では、かなりの蓋然性をもって立証されたのは、「不信または非協力関係は幾何級数的に増幅する傾向にある」ことのみであろう。理論の限界の認識もまた、初学者が早期に心得ておくべき科学的態度であり、指導に留意を要する。

一方、相反するようであるが、想像や類推が科学の領域を広げ、時にはパラダイム次元の飛躍をもたらすことは、数々の天才的碩学の事績が示すところである。自由な仮説の提唱は、仮説の限界の故に排斥される性格のものではない。初学者に指導すべきことは、仮説が十分な立証を経ぬままドグマと化す危険性であって、こうした認識を前提とする限り、伝説はむしろ若々しく大胆であることが推奨されよう。

ゲームの理論ではモデル自体が仮定であり、その所与が現実の社会事象の個別性とは対極の抽象的一般性を帯びているので、自由な連想を許容しやすい面がある。仮説提示の訓練教材としての効用も認め得よう。「囚人のディレンマ」の前記データからの、仮想例を示す。

被験者は二回目のテストで8グループにまとめられた。回答を再記する。

	A	B	C	D	E	F	G	H	計
自白スル	5	4	8	15	15	6	5	9	67
自白シナイ	2	4	2	6	0	2	4	0	20

グループE・Hの特異性が目につく。その対極に、前にもふれたグループB・Gがある。前者をタカ派、後者をハト派と呼ぼう。残余グループは中間派だが、このデータからは、何れの派に収斂するものか、中間派に留まるものか未だ分からない。

さきに与件として、トーナメントによる勝ち上り戦を予定した。かりにアルファベット順に組合わせ、A対B、C対D、……G対Hとして一回戦を行なう。釈放者の率（生き残り率）による順位は下記である⁽⁸⁾。

1位H 2位E 3位A 4位C 5位D 6位B 7位F 8位G

タカ派H・Eの圧勝、ハト派B・Gの惨敗である。二回戦において勝ち上りグループが全部タカ派に転向することは明らかである。したがってこのトーナメント戦の帰結は、

- (1) ハト派全てを含む4グループは、一回戦で敗退。
- (2) 二回戦は全てタカ派同志の勝負。したがって全員不戦敗。

決勝戦すら行なわれないのである。

ところで先に、個々の被験者（囚人）を国家になぞらえたパラフレーズを試みたが、その例よりは、グループを単位とし主権国家と見なすモデル設定が、一層現実に近いであろう。語学のクラスは自発的集団ではないが、成員には他クラスへの対抗意識がある。人は誕生の時に、国家を選べないが、その国家にたいし帰属意識を有している。そしてグループは、クラスの意思決定機関として、国家の政府に照応する。

この国家群は全て、タカ派に非ずば滅亡の選択を迫られるのであろうか。トーナメント戦を国際社会の基本的原理とすれば、そうならざるをえない。冷戦の原理である。ここでは「諸国民の公正と信義に信頼して、安全と生存を保持」する方途はない。

先に挙げた連想より、この世界像がより現実を反映していよう。核漬け対時の閉塞状況にもかかわらず、小国が大戦が不可能であると予想し、あるいはいずれ大戦となれば同じことと観念して、丸腰のまま安全を計るわけにもいかない状況が、このモデルでは提示されている。しかしながら、これとて未だ仮説に至る以前の仮想に過ぎないであろう。国際関係は、現代ではトーナメント戦に類似した構造とばかりは言えない。例えば、国際経済関係はむしろリーグ戦に近い様相を呈する⁽⁹⁾。そこでは、上記8グループの成績評価は別の解析を必要とする。

論者はここで仮説を提示するものではない。ゲームが触発する、豊富な仮説の可能性を指摘したいのである。「囚人のディレンマ」の僅かなテストデータからも、仮説を暗示する連想は湧出する。連想は初学者の討論の中で自然に発生し、そこから仮説の設定まで道は遠くない。

4. 試 行 II

論者は女子学生に対する「囚人のディレンマ」実験の終了にあたり、被験者に課題を与えた。母親に同じ設問を試み、母親の考え方と自分の回答の比較検討を示唆したのである。但しこの場

合、母親自身は囚人ではない。学生が囚人であり、母親が回答を一任されたという設定である。母親は当然、子のため第1の解を選択するであろう。ゲームの体験で得た、協力解への信念がここで再び覆されるであろう。ディレンマは依然として続くのである。それは学生が社会で遭遇し続けるディレンマである。母性本能、家族エゴイズム、社益、国益。これらを圧倒的に是とする環境にあって、それを越える価値の単位を主張し続けることは難しい。社会科学の理論のみならず、実践面での限界も、初学者は早期に覚悟しておくのが望ましい。

課題の回答は回収していない。直接の調査ではなく、正確なデータを期し難いからである。しかし母親、あるいは成人女性のこの面の意識は、論者としても確認の要を感じており、別途調査を行なった。以下、簡単に結果を付記する。

被験対象は、30代から40代の母親50人、都市近郊の団地居住者で、特定の政党支持はなく、地域活動に積極的な層でもない。全員が中流を意識しており、教育への関心は高い。現代日本の平均的母亲とさほど遠くない範囲にあると考えられる。大半は過去に定職についた経験を持つが、現在は不定期のパートタイムのみで、ほぼ専業主婦である。この点ではサンプルの偏りがある。しかし、試行Iの被験者の女子学生の将来像として、比較検討の対象には十分になりうると、考えられる。社会科学の考え方については、学生と同様未経験か、あるいは記憶に留めていない。「囚人のディレンマ」に対する予備知識はない。相互には面識のない10人の母親に、各一人が近所で5人のグループを作り、調査に応ずるよう依頼した。学生のクラスに比し、グループ内の共同体意識は濃密であると考えられるが、反面他グループへの対抗意識は薄い。

テストは10グループ（グループA～J）を集合させ、50人の被験者にたいし同時に行なった。一回目のテストは学生と同一設定、同一方法による、個人単位の回答である。

自白スル 11 自白シナイ 39

第一の解（自然解）の正解率22%は、学生の37%に比し低い。被験者の学歴から推定するに、前記のテストを受けた学生に比し、状況認識や推論能力が相当に劣っているとは考えにくい。論理的理解力が不足したため正解に達しえなかったものと、単純にこの数値を解するのは誤りであろう。被験者は日常、消費者としてあるいは家計の責任者として、自己の損失をいかにして最小にするかの判断については、学生より敏感である。実社会経験は学生より長く、人間性の弱さについての認識も、学生より発達していると見なければならない。第一の解を論理的正解とする判断力は、学生より長じていると考えるのが順当である。ならば被験者が、敢えて第一の解の選択を避けた理由は何であろうか。円満な社会人に成長する過程で、第一の解を越えて第二の解の認識に達していると想定すべきか。もしそうであれば、被験者は生活体験を通し、自然に社会科学

の理念を体得したといえよう。自然に体得しうるものであれば、教育の場における社会科学の理念や手段の学習の重要性を、さほど声高に説くことはなかろう。

被験者が、第二の解（協力解）の認識段階に達しているとするのは、皮相的観察である。被験者は、先に検討したプロセスを踏んで、第二の解の認識に達したのではない。彼らの選択の動機は、自我を越えた価値単位の拡大の論理ではない。二回目、三回目のテスト結果を検討すると、それは明らかになる。

二回目のテストは、設定を一部変えて行なった。「囚人」は被験者本人ではなく、自分の子供と変更したのである。母親に回答が委任されていると仮定する。一回目同様に、個々人が単独で考慮、判断した回答を、無記名で回収した。

自白スル 24 自白シナイ 26

第一の解への急上昇が見られる。

三回目の回答は、グループ単位の討論の後、回収された。グループは近所同志の5人であり、「囚人」である団地棟の住民を代表する。討論方式など他の条件は学生と同一である。

	A	B	C	D	E	F	G	H	I	J	計
自白スル	0	0	0	0	0	1	0	0	1	0	2
自白シナイ	5	5	5	5	5	4	5	5	4	5	48

ここでは第一の解は殆ど忌避されている。二回目と三回目の回答を比較すると、相反する二つの傾向が観察される。二回目では第一の解へ、三回目では第二の解へ、数値は大きく動く。

学生の場合、討論後の回答では、第一の解へ向かう数値変化が見られ、これは責任意識と討論による情報が相まって、第一の解の論理性への、被験者の客観的認識が深められたものと解析した。母親の回答は、2回目テストでは学生と同じ傾向を示す。子に対する責任感が、第一の解へ向かったと見られる⁽¹⁰⁾。しかし、討論の効果については、学生と正反対の結果をもたらした。しかも数値は、過激なまでの変化を示す。これは何を意味するか。

討論によって一挙に、第二の解へ向かう、価値の視点の革命が起こったとは解し難い。母親たちの場合、学生が討論を介して得た認識の段階的变化は、実は生じていないと考えるべきではないか。母親の選択を促すのは、論理ではなく、別の要因である。論理であれば、二回目三回目のテストに、このように極端に相反するベクトルが観察される可能性は少ない。

学生は、討論による「囚人のディレンマ」体験を通し、自己と他者が不分明な世界から、先ず

中心価値としての自己を発見し、それに対峙する存在としての外界＝他者を分離した。次いで、自己の価値実現の方策として、他者との価値共有を発見するのである。これらの発見に照応する社会科学の方法や理念の局面については前述した。

母親たちの世界は終始、自己と他者が不分明なのである。一回目から三回目まで、認識は一貫しているのであり、数値の変化は意識の変化を意味しない。この被験者においては、価値基準はこの不分明な世界に所与として存在する。常にその基準に従った転移を行なったにすぎない。

一回目のテストでは被験者は、協力解を選択した。それは所与たる価値としての「道德」に従った選択にすぎない。二回目のテストでは「道德」の解釈の混乱が数値に現われた。「無償の母性愛」と「博愛」の選択に迷ったのである。三回目、討論により公定の「道德」の確認が行なわれ、ほぼ全員がそれに従った⁽¹¹⁾。二回目のテストでは、5人中3人が「自白スル」を選択したグループが3組あったが、このグループはいずれも、三回目テストでは全員が「自白シナイ」を選択した。選択基準は個々の信念ではないので、多数決原理は機能しない。「何が道德か」で定められる。より正確には「何が道德とされているのか」である。言換えれば、「価値基準」や「道德」は、ここで「建前」と同義である。それはまた、容易に「社益」や「国益」に読み替えうる。このテストでは一斉に「自白シナイ」を選択し、「博愛」に忠誠を誓った被験者は、「建前」の側の事情変更にともない、一斉に「自白スル」可能性をひめている⁽¹²⁾。

さきに社会科学の機能の一つとして、主体的、科学的に、目的意識をもって社会事象に正対する人間の育成を挙げた。上記の被験者の状況は、こうした人間像から、絶望的に遠い。この国においては、実社会に身を置くことのみで、より良き社会を目指す方法と理念を、人間が自然に体得することは、難しいと思われる。教育現場において、初学者に社会科学的思考の訓練を施す意義は、未だ大きいと言わねばなるまい。

本稿では、「囚人のディレンマ」のゲーム体験を通じて、段階的に獲得される論理的認識について若干の考察を加え、社会科学の入門指導におけるその応用を論じた。ゲーム実施に際して、初学者間での討論の併用を、認識の獲得に有効な具体的手段として提唱したものである。

試行の評価分析に当たって、副次的に国際関係などへの論及を行なったが、本稿のデータのみでは、この論及部分は仮説を提示する域に至っていない。仮説を暗示したに留まる。

注

(1) 本稿の例の「死刑あるいは釈放」の組合わせの他、「刑期10年、自白すれば2年減刑」のように、穏健な設定もある。前者は国際政治のモデル、後者は国際経済のモデルに妥当しよう。論者の場合、原理的には両立しない価値体系の共存を、検討するためのモデルとして使用した。

(2) Minimax 原理による解。ゲーム理論の創始者、J. von Neumann は、マトリックスによる証明を行な

っているが、論者は受講者に分かり易く説明するため、下記数式によった。

死刑 = A 懲役10年 = B 釈放 = C 賞金付釈放 = D 価値評価は $A < B < C < D$

相手が自白する確率 = x 相手が自白しない確率 = y $0 \leq x, y \leq 1$ $x + y = 1$ 自分が自白する場合の期待値 = M 自分が自白しない場合の期待値 = N とする

$M = Bx + Dy$ $N = Ax + Cy$, $Bx \geq Ax$ かつ $Dy \geq Cy$ でありまた x と y が同時に 0 の場合はないから、 $Bx + Dy > Ax + Cy$ 即ち $M > N$

A, B, C, D に特定の数値（例えば死刑 = -100 など）を与える例もあるが、定量評価は主観によって異なるので、社会事象のモデルとしてゲーム理論を応用する場合、かえって minimax 原理の論理性に余計な誤解を与えると考える。尚上記の数式、また以下の注の計算式も、女子学生には負担が大きく、実際の講義では省略したほうが、本稿の意図にかなうと思われる。

- (3) 厳密に言えば、このモデルでは、 $A + D < 2C$ は当然には前提されない。つまり賞金がそれだけで一方の死を犠牲にするほど、評価される場合の問題がある。しかし実際問題として、医療保険給付の限度額など、特殊な問題を扱わない限り、差し支えないだろう。
- (4) A. Rapoport と A. Chammah (University of Michigan) の調査による。被験者の国籍により、実験結果の数値は異なるであろうし、数値の違い自体が研究対象になりうると思われる。本稿の分析の基準値としては、日本人に対する組織的実験結果のデータが望ましいのであるが、論者には入手不能であった。しかしながら、論旨の主眼は、一回目テストでの自然解の正解率を分析することにはなく、自然解から協力解にいたる、論理的思考に関するものである。この場合基準値の大きさではなく、増減を考えれば十分であり、アメリカの実験結果を使用しても、変りないものとする。実際、日本の大学において、「囚人のディレンマ」を論ずる際にはこの数値をそのまま使用している。
- (5) 国際政治や市場理論専攻の学生は、ゲームの予備知識があり、実験者の意図に添った回答を行なうので、実験の意味はなお失われる。
- (6) 例えばトーナメントのAブロックとBブロックで、Aブロックでは1グループが、Bブロックでは2グループが、タカ派であったとする。Aブロックでは、タカ派の1グループが勝ち上がるが、Bブロックではタカ派は2グループが対戦した時点で不戦敗となる。最終勝者はしたがってAブロックのタカ派グループである。
- (7) 比率は確率計算により求めた。

テスト総人数 = N 「自白スル」とした回答者数 = P 「自白シナイ」とした回答者数 = Q
とする。被験者のうち、任意の二人 (A, B) を選び、「囚人」とする。

A, B が両者とも釈放されるのは、A, B ともに「自白シナイ」場合である。

A が「自白シナイ」とした回答者である確率は Q/N

A が「自白シナイ」回答者であるとして、B が「自白シナイ」回答者である確率は、
 $(Q-1)/(N-1)$

A, B が両者とも「自白シナイ」とした回答者である確率は、したがって、
 $Q/N \times (Q-1)/(N-1)$

被験者全員を、一斉に二人ずつ任意に組み合わせたととき、どの組にたいしても、上記の確率は等しく該当する。

A, B が両者とも懲役になるのは、A, B ともに「自白スル」場合である。この確率は同様に
 $P/N \times (P-1)/(N-1)$

残る一つが、一方死刑一方釈放の場合であり、以上三つの場合以外の結果はありえないので、A, B の一方が死刑、一方が釈放になる確率は

$$1 - \{Q/N \times (Q-1)/(N-1) + P/N \times (P-1)/(N-1)\}$$

この計算式により得られた、各場合の確率を、パーセンテージ比率表示とした数値である。87人の奇数

では、実際に組合わせた場合、端数が生じるが、確率計算上は無視してよい。

確率であるからには、実際の個々の組合せ結果とは、相当の数値の違いがありうる。極端な場合、例えば全員が「自白スル」者同志、「自白シナイ」者同志の組合せになったとすれば、「死刑」は一人もいない。確率計算による比率は、組合わせた無限回数行なった結果、それを平均した比率に等しい。一種の理論的極限值である。

- (8) 全グループを同人数にした補正值により、釈放者の率を計算した結果である。

国際関係のモデルとして考えるなら、人数にばらつきがあるほうが、現実に近いのだが、そうすると、大国、小国で別の戦略が生じてくる。それも興味深い実験となるであろう。このテストでは、数値補正を予告しており、各グループの判断に人数のハンディは意識されていない。釈放者の率は次のように計算される。

Aグループの人数 = A_n Aグループの内「自白スル」を回答した人数 = A_p

Aグループの内「自白シナイ」を回答した人数 = A_q

Bグループ以下についても同様に、 $B_n, B_p, B_q, \dots, H_n, H_p, H_q$ とする。

Aグループの任意の一人がBグループの任意の一人と「共犯者」として組み合わせられた場合、彼が釈放されるのは、彼がBグループの内「自白シナイ」者と組み合わせられた場合であるから、彼の釈放の確率は、 B_q/B_n である。この確率はAグループ全員について等しいので、Aグループが期待しうる釈放者数の期待値は、 $A_n \times B_q/B_n$ となる。

同様に、BグループがAグループと組み合わせられた場合、Bグループの釈放者数の期待値は

$B_n \times A_q/A_n$ となる。

ちなみに、補正值を A_n', B_n', \dots, H_n' とすると、

$A_n' = B_n' = C_n' = \dots = H_n'$ と設定し、

$A_q/A_n = A_q'/A_n' \quad B_q/B_n = B_q'/B_n' \dots H_q/H_n = H_q'/H_n'$

- (9) 国際経済関係のゲームとして、与件に定量値を設定し、勝負ではなく複数のグループの達成度を競うモデルが考えられる。ハト派同志のリーグ戦、タカ派リーグ戦などを、ブロック経済のアナロジーとして検討する例など。このモデルでは、協力解の発見は容易である。

- (10) 子は母にとって自我の延長に過ぎず、他者への責任感とは異なる、とする反論もあるが、この場合は当たらない。責任感とはここでは、己以外の視点に立ってみることに過ぎない。「自分の共犯者なら信じたい、子の共犯者は信じられない。」という感情には、前段では希望の投影としての主観的世界像があり、後段ではそれと対称をなすように客観的な認識がある。それは子という「他者」を媒体として生じたものである。

- (11) グループ討論の内容を聴取するに、討論では、所属グループの釈放の可能性をいかに高めるかの観点からの検討が殆ど行なわれていない。理論的にそれを考えていなかったということではない。「口に出しにくかった」のである。「誰かが言い出せば、賛成したであろうが、自分が言い出して、利己主義な人と思われるのが、嫌だった」結果、本筋の検討が討論に上っていない。団地棟全員のためという大義名分も、「意見なんてどう言う風に伝わるか分からない」苦い思いを実生活で体験している母親には、頼りになっていない。討論はおもに、二回目のテストでどう回答したかの、「意見交換（探りあい）」に始まり、いささかの論拠表明もあったが、「人の子も自分の子も同じ人間」の類の発言が一旦あると、「やはりそうしなければね。難しいけど」と留保付きながら、同意しなければならないのは、「お互いの子供も知ってるし」本題からはずれた配慮を優先しなければならないからである。

一方建前発言の張本人とても、自分の発言を信念としているわけではない。「そういわれて感心したことがあるから、そういっておけば間違いないと思って」発言したにすぎない。「一人だけ偉そうにと、思われたら」何時でも発言は撤回し多数意見に従うつもりであるという。

- (12) 母親に対する調査は、親密度の高い自発的グループを単位として行なわれたので、学生との数値のみを

単純に比較すべきではない。注(10)に見られるように、自分の本心を周辺の人々に知られることを警戒し、実際よりは「建前」回答の傾向が強いデータといえよう。しかし、それを考慮に入れたうえ尚、ここで分析した被験者の心象は、日本の大多数の成人女性の考え方の一面を現わすものといえよう。逆に言えば、学生の大多数が10年後、再びこうした心象に舞い戻らないという、保証はない。