

第 6 部

関連研究

＜非＞寛容の限界

－T.シェリングのモデルをもちいた分析－

数土 直紀
(学習院大学)

【要旨】

本論では、協同的に達成される秩序とは異なった差別的な秩序の形成について議論する。また本論では、差別的な秩序の形成を説明するモデルとして T.シェリングのモデル (Schelling 1978) をもちいてシミュレーションをおこなった。その結果、差別的秩序をもたらす異集団への非寛容さにも限界があること、また二集団のケースと三集団のケースとでは秩序形成のメカニズムが異なることが明らかにされた。

キーワード: 差別的な秩序, シェリングのモデル, シミュレーション

1. はじめに

秩序という術語で想起されるイメージは、社会科学においては一般的に協同的な秩序であることが多い。例えば、社会的ジレンマを解決することによって実現されるような秩序とは、まさにこのような協同的な秩序だといっているだろう。個人はこの協同的な秩序に組み込まれることでかえって利得を増大させることが可能になるので、このような秩序を望ましいものとして積極的に評価することができる。しかし、すべての秩序がこのような望ましい性質を備えているわけではない。たとえば、集団が個人の（社会的）属性によって階層化されているような、そんな差別的な秩序も存在する。このように個人を差別的に配置し、またそうした秩序にしたがって資源を不均等に配分するようなことがあれば、そうした差別的な扱いを帰結する秩序は一般的には望ましくないと考えることができる。

このような差別的な秩序がどのようにして形成されるのかという問題に対して、T. シェリングのシミュレーションをもちいた分析 (Schelling 1978) は興味深い知見を明らかにしている。シェリングは、自分と異なる属性をもった個人を排除しようとする傾向をもたなくても、自分がマイノリティになることを避けようとする欲求をもっていることを仮定するだけで、居住区域の分離が発生することを明らかにしているからである。シェリングが分析にもちいたモデルは、ゲーム理論における囚人のジレンマと同様に、きわめてシンプルなものであり、したがって明らかにされた事実の含意はきわめて応用範囲の広いものと思われる。にもかかわらず、これまでシェリングのモデルを発展させた研究は多くないことが複数の研究者によって指摘されている (Zhan 2004, 籠谷 2004)。

そこで本論では、やや条件を変更した形でシェリングのモデルを追試し、その結果の頑健性を確認することを目的とする。それと同時に、集団が多様性をもつことが差別的な秩序の形成に対してどのような影響を与えうるのかを考察する。

2. 基本モデル

2.1 モデルの仮定

それでは、本論でシミュレーションに用いるモデルがどのようなものであるかを明らかにしよう⁽¹⁾。本論でもちいられるシミュレーションの基本的な仮定は、以下の通りになる。

まず、世界は、10行×10列の100の升目に区切られている。各升目には一個体だけが居住することができる（しかし、すべての升目に個体が居住している必要はない）。そして、どの升目にどの属性をもつ個体が居住することになるか（あるいは、居住者のいない空の升目になるか）は、パラメータとして定められた確率にしたがって、ランダムに決定される。

つぎに、各個体はそれぞれに周囲が自分と異なった属性をもった個体に囲まれることをどの程度まで許容することができるか、分離度に対して許容水準をもつものとする。このとき各個体は、自分が居住する周辺地域に、（自身を含め）自分と同じ属性をもつ個体の居住している割合が与えられた許容水準の数字を下回ったとき、他の場所へ移動しようとする。もし許容水準を上回る空き地を探し出すことに成功すれば、移動を動機づけられている個体はその空き地に移動する⁽²⁾。しかし、もし許容水準を上回る空き地を見つけ出すことができなければ、移動をあきらめるものとする。とうぜん、許容水準は自分の移動によってだけでなく、他の個体の移動にもよって変化するので、移動のチャンスは複数回与えられる。本モデルでは、すべての個体にあらかじめ定められた回数の移動のチャンスを与えられる。

ここで、移動前と移動後の全体の居住分布がどのように変化するかを確認しよう。

例えば、図1が移動前の状況になる。図1では、○が0.33の確率で、また●が0.33の確率で居住するようにパラメータが設定されている。確率でランダムに決められるため、○と●の割合は必ずしも同じになっていない。

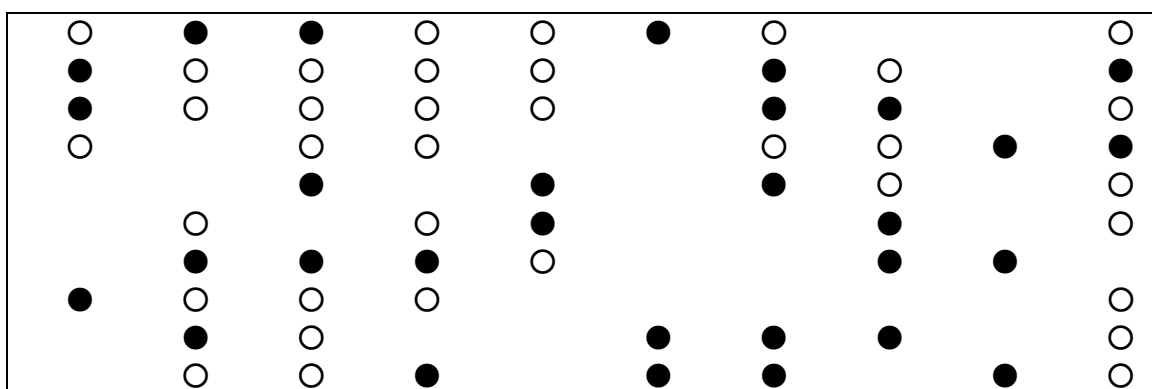


図1 移動前の状況

このとき、○と●にそれぞれ0.45という許容水準を与え、30回の移動のチャンスを与えると居住分布は図2のように変化する。

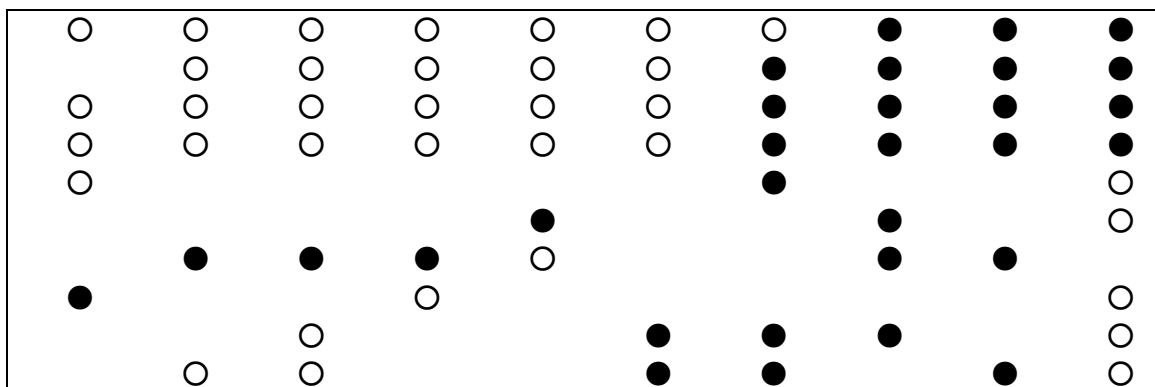


図2 移動後の状況

図2からわかることは、最初は混住していた○と●がはっきりと分離されたということである。0.45という許容水準は、自分もふくめて同じ属性をもった個体が周囲に半分以下でも「移動しなくてもよい」と判断しているのだから比較的緩やかな条件とっていいだろう（さらに、9つある升目のうち一つは自分が占めているので、0.45という条件は（自分をカウントしない場合と比較すると）容易に達成される）。しかし、このような許容水準の下でも、このようなはっきりした分離が確認される。これは、いわばシェリングが明らかにした知見の正しさを裏付けるものといっていいだろう。

2.2 分離度の計算

図1と図2を比較すると、○と●は移動を繰り返すにことよって分離されていることが確かにわかる。しかし、○と●の分離を判断するとき、このような理解の仕方はいわば直観に頼ったものでしかないことに注意する必要がある。したがって、○と●が移動を繰り返すことよって分離したという主張は、この段階では決して厳密なものとはいえないだろう。このような難点を克服するためには、それぞれの居住のパターンに応じて、○と●がどの程度分離されているのかを、厳密に測定する必要があるだろう。そして、そのような測定を経てはじめて、パラメータの変化によつて分離の程度がどのように変化するかを明らかにすることができる。

したがって本論考では、移動前と移動後で分離の程度がどのように変化するか、あるいはパラメータの変化によつて分離の程度がどう変化するか、このことを正確に判断するために分離度を計算することにしよう。

本論考では、分離度を次のように計算する。

まず、ある個体について、その個体が居住する升目とそれに隣接する合計9つの升目にいくつの個体が居住しているかを計算する。このとき、計算には、その個体自身もカウントされるとする。

次に、カウントされた個体の中で自分と同じ属性をもった個体の数（その個体が○であれば○の数、その個体が●であれば●の数）を計算し、その数を先にカウントした個体数で除する。このとき値は、自分が居住する周辺地域に、（自身を含め）自分と同じ属性をもつ個

体がどの程度居住しているか、その割合を示すことになる。

最後に、すべての個体について同様の計算を行い、その平均値を計算する。そしてこの値が、全体が○と●にどの程度分離されているかを示す分離度となる。この数値は、同じ属性をもった個体数に周囲が囲まれている割合が高ければ高いほど、つまり分離の程度が高ければ高いほど高くなるからである。

実際に、図 1 と図 2 でそれぞれの分離度を計算することにしよう。図 1 では、分離度は 0.611 となる。図 2 では、分離度は 0.876 となる。確かに、図 1 の場合と比較して、図 2 の方が分離の程度が激しいことがわかる。

3. 分析結果

3.1 二集団の場合

3.1.1 <非>寛容の限界

本節では、シミュレーションによる分析結果を明らかにしていく。

最初に、許容水準をパラメータにして、許容水準の変化によって分離度がどのように変化するのかわ確認した。ここでは、すべての個体について 30 回の移動機会を与えた試行を各々の許容水準について 30 回繰り返し、そうして得られた分離度の平均を図 3 によって示した。

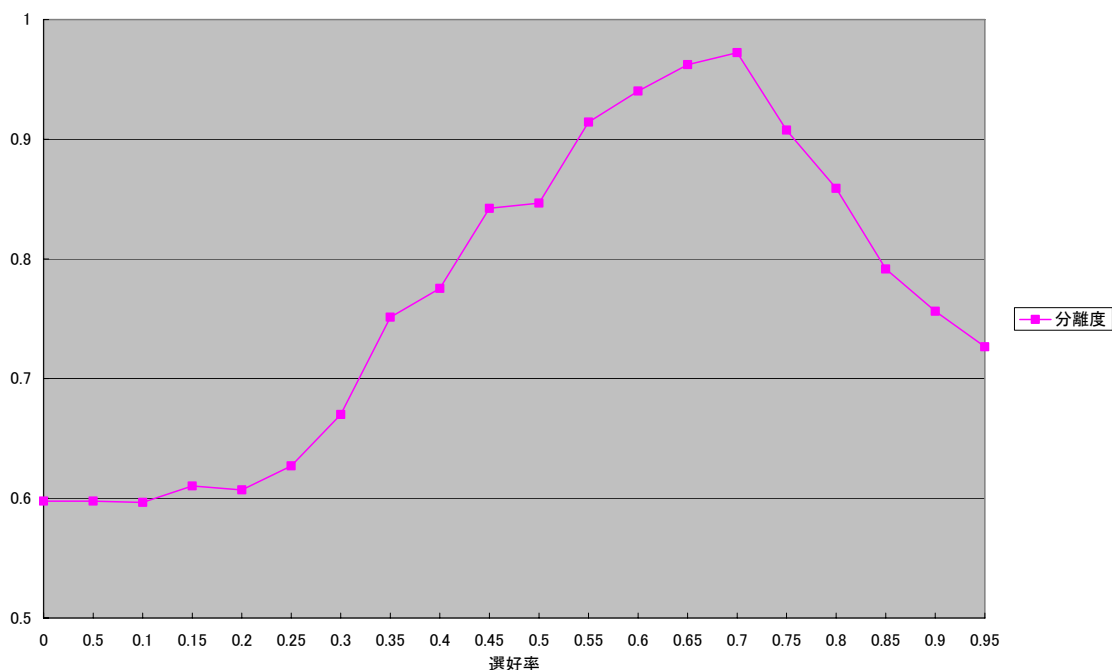


図3 許容水準(選好率)による分離度の変化

また、一元配置の分散分析を行い、多重比較の結果を参考にして許容水準を3つのグループに分類した。なお、図3をみればほぼ自明であるが、許容水準によって分離度の平均に関

して統計的に有意な差を確認することができる ($F(19,580)=92.465, p<0.01$) .

表1 許容水準によるグループ分け

グループ	許容水準
分離がほとんど生じない	0.05 0.10 0.15 0.20 0.25
分離が生じる	0.30 0.35 0.40 0.45 0.50 0.80 0.85 0.90 0.95
強い分離が生じる	0.55 0.60 0.65 0.70 0.75

シェリングが明らかにした重要な知見は、個体が比較的寛容な許容水準をもっている、寛容な許容水準から予想することの難しい、はっきりとした分離が生じるということである。移動がないときの平均的な分離度が 0.6 前後なので、0.6 よりも低い許容水準でも分離が生じていることからシェリングの主張の正しさを確認することができる。しかし同時に、許容水準が 0.25 を下回るような場合には分離がほとんど生じないことから、そのような分離がつねに起こりうるものでないこともわかる。やはり、すべての個体が非常に寛容なときは、分離は発生しないのである。

興味深いのは、分離を生じさせるという点について、非寛容にも限界があったという事実である。具体的には、許容水準が 0.7 を越えた後は、その後はいくら許容水準を高めにとっても分離度は上らず、むしろ減少するということである。このような現象が生じるのは、個体があまりにも高い許容水準をもつと、条件を満足する空き地を発見できなくなるためにかえって移動できないケースが発生するからである。したがって、許容水準と分離度の関係は線形ではない。許容水準が変化するとき、激しい分離を引起す範囲が存在する。仮に個体もっている許容水準が分離度を規定しているにしても、その関係は単純でないことに注意する必要がある。

3.1.2 集団サイズの影響

シェリングが明らかにした知見の一つに、二集団の大きさに差がない場合よりも、差がある場合の方が激しい分離が生じるというものがある。もしこの知見が正しいのだとすると、分離度は個体もつ許容水準に影響をうけるだけでなく、二集団の大きさの差にも影響をうけることになる。そこで、シェリングの明らかにしたこの知見が、本論で使用しているモデルについても妥当するかどうかを検討してみた。

許容水準を 0.45 に、1 試行につき移動機会を 30 回に固定し、30 回の試行を行った。このとき、分離度の平均は以下ようになった。上段が両集団のサイズに大きな差が出ないようにしたケースであり、下段が両集団のサイズに大きな差が出るようにしたケースである（具体的には、前者は○と●がそれぞれ 0.33 の確率で、ある升目に居住するように制御し、後者は○が 0.15 の確率で、●が 0.50 の確率で、ある升目に居住するように制御した）。

集団サイズに差がない場合 移動前 平均分離度=0.60
移動後 平均分離度=0.83

集団サイズに差がある場合 移動前 平均分離度=0.72
移動後 平均分離度=0.88

それぞれの結果をみても、移動前と移動後ではいずれの場合も分離度が大きく変化していることがわかる。また、集団サイズに差がある場合と差がない場合を比較すると、後者のケースの移動後の平均分離度が前者のケースの移動後の平均分離度を確かに上回っており、しかもその差は統計的に有意である($t=-4.921$, $p<0.01$)。したがって、本モデルにおいてもシェリングの知見の正しさを確認することができる。

しかし、このとき注意しなければならないことは、そもそも両者には移動前から平均分離度の大きな差があるということである。そして、この差は統計的にも有意である($t=-11.98$, $p<0.01$)。したがって、両者の差は移動によって生じたのではなく、移動前にすでに存在していたのであり、さらにいうならば移動することによって両者の差はむしろ縮まっている。

では、なぜこのような平均分離度の差が生じるのだろうか。このメカニズムを明らかにするために、集団別に平均分離度がどのように変化しているのかを確認することにしよう。

集団“○”について

集団サイズに差がない場合 移動前 平均分離度=0.59
移動後 平均分離度=0.83

集団サイズに差がある場合 移動前 平均分離度=0.39
移動後 平均分離度=0.72

集団“●”について

集団サイズに差がない場合 移動前 平均分離度=0.59
移動後 平均分離度=0.83

集団サイズに差がある場合 移動前 平均分離度=0.82
移動後 平均分離度=0.92

集団サイズに差がない場合には、集団に関係なく平均分離度の移動前も移動後も数値は似たようなものになる（コンマ3桁以下を四捨五入しているため、同じ数値になっているが厳密には異なった数値である）。しかし、集団サイズに差がある場合には、移動前と移動後では大きく異なったものになっている。個体数の大きい●の移動前の平均分離度は、同じ属性をもった●にであらう確率の高さを反映して、きわめて高いものになっている（逆に、○の移動前の平均分離度はきわめて低いものになっている）。このことから、個体数の大きい●の平均分離度の高くなっているために全体の平均分離度が引き上げられていることがわかる。

また、移動前と移動後の平均分離度の変化の度合いを比較すると、○の方が●よりも変化の度合いが大きいことがわかる（ $0.72-0.39=0.33$, $0.92-0.82=0.10$ ）。したがって、激しい分離

は、ばらばらに居住していた、マイノリティである○が集住しはじめることによって発生する部分が大きいこともわかる。

3.1.3 許容水準の影響

前項では、集団のサイズに差がある場合、許容水準による移動によって生じる分離の程度にどういった差があるかを検討した。その結果、基本的にはシェリングが確認した事実を再確認することができたが、同時にそのような差が生じるメカニズムについて検討した。

本項では、集団のサイズにではなく、集団によって許容水準が異なる場合、許容水準による移動によって生じる分離にどのような違いが現れるかを検討したい。

本論では、ある升目に○が居住する確率もしくは●が居住する確率をそれぞれ 0.33 に固定し、1 試行につき移動機会を 30 回に固定し、30 回の試行を行った。このとき、分離度の平均は以下ようになった。上段と中段が両集団の許容水準に大きな差が出ないようにしたケースであり、下段が両集団の許容水準に大きな差が出るようにしたケースである。具体的には、上段では○と●の許容水準を 0.35 に揃え、中段では○と●の許容水準を 0.75 に揃えてシミュレーションを行ったときの平均分離度である。また、下段は○の許容水準を 0.35 に、●の許容水準を 0.75 に固定した場合の平均分離度である。

< 1 > 許容水準が異なる場合 (ともに 0.35)	移動後の平均分離度 = 0.75
< 2 > 許容水準が異なる場合 (ともに 0.75)	移動後の平均分離度 = 0.91
< 3 > 許容水準が異なる場合 (○=0.35, ●=0.75)	移動後の平均分離度 = 0.81

結果をみると、許容水準が異なる場合の平均分離度は、許容水準が 0.35 であるケースとも、また許容水準が 0.75 であるケースとも異なっていることがわかる。許容水準が異なる場合、移動後の平均分離度は、許容水準が高い場合の平均分離度と許容水準が低い場合の平均分離度の中間に近い値をとっている。なお、各々の平均分離度の差はいずれも統計的に有意な値を示しており (< 1 > と < 2 > : $t=-6.372$, $df=29$, $p<0.01$ < 1 > と < 3 > : $t=4.31$, $df=29$, $p<0.01$ < 2 > と < 3 > : $t=-3.67$, $df=29$, $p<0.01$)、少なくとも本論で設定した条件の下では繰り返し生起することが予想される。

したがって、許容水準が異なる場合、分離形成のメカニズムはいずれかの集団の許容水準に引きずられるのではなく、ちょうど両者の平均程度 of 分離を引起すことがわかる。このことは、異集団に対して敵対的な許容水準をもつ集団によってもたらされる分離・集住への圧力を比較的寛容な許容水準で対応することで和らげることができると解釈できる一方で、逆に異集団に対して比較的寛容な許容水準をもっているにもかかわらず敵対的な許容水準をもっている集団と対面している場合には、結局や分離・集住の形成に寄与してしまわざるをえないとも解釈できる。分離は相手との相互作用によってもたらされる現象であり、仮に自身が比較的寛容な許容水準をもっていたとしても、それだけでは分離の形成を完全に妨げることは難しいといえるかもしれない。

全体としての分離の程度を問題にしたが、各集団内ではそれぞれどのような変化が生じて

いるのだろうか。許容水準の違いがもたらす問題をさらに詳しく見るために、集団別に平均移動率がどのようになっているのかを確認することにしよう。

集団“○”について

< 1 > 許容水準が異なる場合 (ともに 0.35)	移動後の平均分離度 = 0.762
< 2 > 許容水準が異なる場合 (ともに 0.75)	移動後の平均分離度 = 0.935
< 3 > 許容水準が異なる場合 (○=0.35, ●=0.75)	移動後の平均分離度 = 0.809

集団“●”について

< 1 > 許容水準が異なる場合 (ともに 0.35)	移動後の平均分離度 = 0.732
< 2 > 許容水準が異なる場合 (ともに 0.75)	移動後の平均分離度 = 0.929
< 3 > 許容水準が異なる場合 (○=0.35, ●=0.75)	移動後の平均分離度 = 0.811

各集団の許容水準が異なる場合には、集団に関係なく平均分離度の移動前も移動後も数値は似たようなものになる。移動回数の少ない許容水準が 0.35 のケースでは、集団間でやや分離度に差があるが、とうぜん統計的には有意な差ではない。つまり、許容水準に大きな差があるにもかかわらず、移動によってもたらされる分離度については両集団に差はなくなってしまうのである。これは、全体の分離度を問題にしていたときと同様に、仮に比較的寛容な許容水準を持っていたとしても差別的な集団と居住区域を共有するときには分離・集住のプロセスに巻き込まれてしまう。しかし分離・集住の程度は、差別的な集団同士の場合よりは比較的緩やかなものになり、この点は集団に関係なくあてはまる。

3.2 三集団の場合

3.2.1 集団数の影響

前項では、想定される集団の数は○と●の2つであった。したがって、ある集団に所属する個体にとって、自身と異なる属性をもった個体は1タイプしか存在しなかった。そして、このように二集団しか存在しないようなケースでは、自分がマイノリティになるような地域に居住し続けることは避けたいという欲求を個体もっているとは仮定すると、その欲求が比較的弱いものであったとしても、結果として激しい分離・集住の生じる可能性があることが明らかにされた。

それでは、想定される集団の数を2から3とか、4とか、あるいはそれ以上に増やした場合にも、激しい分離・集住をともなった差別的な秩序が形成されるのだろうか。仮にそのようなケースでも差別的な秩序が形成されるとして、それは二集団の場合のときと同じような結果をもたらすのだろうか。

この問いは、いわば社会における多様性の役割への問いでもある。ここでは、社会にさまざまなタイプの集団が混住しているとき、そこで展開される移動・集住のパターンが二集団の場合と同じであるかないかを問題にしていこうとする。

まず、モデルの設定を確認することにしよう。前節までは、○と●といった2タイプの個体のみを想定してきたが、ここではさらに△というタイプの個体を想定する。そして、すべての升目は、それぞれ確率 0.25 で○か、●か、△が居住することになる。そうやって、すべての個体の最初の居住地が決まったあと、各個体はあらかじめ与えられた許容水準にしたがって、移動するかしないかの判断を下す。

ここでは、すべての個体に移動機会を 30 回与える試行を 30 回行い、試行ごとに求められた分離度を平均した値を比較するものとする。最初に比較されるのは、許容水準を 0.50 に設定したケースと、0.30 に設定したに設定したケースである。

三集団の場合

0.50 に設定した場合 移動前 平均分離度=0.44
移動後 平均分離度=0.83
移動後－移動前 0.83-0.44=0.39

0.30 に設定した場合 移動前 平均分離度=0.45
移動後 平均分離度=0.65
移動後－移動前 0.65-0.45=0.20

上述の数字は三集団のケースであるが、比較のために二集団（発生の確率はそれぞれ 0.33）のケースについても結果を示しておこう。

二集団の場合

0.50 に設定した場合 移動前 平均分離度=0.61
移動後 平均分離度=0.86
移動後－移動前 0.86-0.61=0.25

0.30 に設定した場合 移動前 平均分離度=0.62
移動後 平均分離度=0.71
移動後－移動前 0.71-0.62=0.09

この結果から、二集団から三集団になってもやはり分離・集住がおきることがわかる。そして、移動後の分離度と移動前の分離度との差が、分離圧力の強さを示していると考えられるならば、むしろ集団の数が増えたケースの方が強まっていることがわかる。

3.2.2 集団サイズの影響

前項では、勢力が拮抗している場合には集団の数が増えることで分離・集住への変化の度合いが増すことが確認された。それでは、集団のサイズに差がある場合にも同じような結果を確認することができるのだろうか。二集団の場合とは異なり、三集団の場合には集団のサ

サイズに差があるケースを2つに分類して考えることにする。1つは三集団の中で圧倒的なマイノリティが存在するケースである。具体的には、三集団のうちの二集団はサイズが対等か、少なくとも大きな差がないのに対して、残りの一集団はサイズがきわめて小さくなっているようなケースである。もう1つは、三集団の中に圧倒的なマジョリティが存在するケースである。具体的には、三集団のうちの二集団はサイズが対等か、少なくとも大きな差がないのに対して、残りの一集団のサイズがきわめて大きくなっているようなケースである。

ここでは集団サイズをかえて個体の移動によって分離度がどのように変化するかをみてみることにしよう。前者のケースを検討するために、○を確率 0.35 で、●を確率 0.35 で、△を確率 0.10 で発生させる。この場合、△がマイノリティになる。後者のケースを検討するために、○を確率 0.5 で、●を確率 0.10 で、△を確率 0.10 で発生させる。この場合、がマジョリティになる。いずれも、許容水準は 0.40、移動機会は 30 回に設定し、30 回の試行によって得られた平均を計算している。

まず、圧倒的なマイノリティが存在するケースでの分離度の変化は次のようになった。

全体	移動前	平均分離度 = 0.50
	移動後	平均分離度 = 0.73
“○”	移動前	平均分離度 = 0.54
	移動後	平均分離度 = 0.77
“●”	移動前	平均分離度 = 0.53
	移動後	平均分離度 = 0.76
“△” (少数派)	移動前	平均分離度 = 0.25
	移動後	平均分離度 = 0.50

次に、圧倒的なマジョリティが存在するケースでの分離度の変化は次のようになった。

全体	移動前	平均分離度 = 0.64
	移動後	平均分離度 = 0.77
“○” (多数派)	移動前	平均分離度 = 0.77
	移動後	平均分離度 = 0.85
“●”	移動前	平均分離度 = 0.30
	移動後	平均分離度 = 0.53
“△”	移動前	平均分離度 = 0.29

移動後 平均分離度=0.57

いずれの結果も、移動によってやはり分離・集住が発生していることがわかる。しかし、発生メカニズムは二集団の場合とは異なっている。二集団の場合ではサイズの小さな集団に所属する個体が集住することで分離が発生していたが、三集団の場合ではいずれもサイズの小さな集団の平均分離度はサイズの大きな集団の平均分離度を下回っている。むしろ平均分離度の数値の変化からは、サイズの大きな集団がサイズの小さな集団に属する個体を排除する形で分離・集住が強化されているように解釈できる。

3.2.3 許容水準の影響

前項では、集団のサイズに差がある場合、分離・集住の発生するメカニズムが二集団のケースと三集団のケースとでは異なってくる可能性のあることが明らかにされた。しかしここでは、三集団の間で許容水準は異なるものと仮定していた。もし集団間に許容水準に差があるとき、そのことは分離・集住のメカニズムにどのような影響を与えるのだろうか。すでに二集団のケースについては前節で検討しているけれども、本項では三集団についても検討することにしたい。

ここでは、三集団の場合には許容水準に違いがあるケースをやはり2つに分類して考えることにしよう。1つは三集団の中で極端に寛容的な集団が存在するケースである。具体的には、三集団のうち二集団は許容水準が同程度であるのに対して、残り一集団の許容水準がとても低い数値（つまり、より寛容的）に設定されているようなケースである。もう1つは、三集団の中に極端に非寛容的な集団が存在するケースである。具体的には、三集団のうち二集団は許容水準が同程度であるのに対して、残り一集団の許容水準がとても高い数値（つまり、より非寛容的）に設定されているようなケースである。

ここでは許容水準をパラメータにして、分離度がどのように変化がするのかをみてみることにしよう。前者のケースを検討するために、○と●の許容水準を0.50に設定し、△の許容水準を0.1に設定する。この場合、△がとりわけ寛容的な集団となる。後者のケースを検討するために、○の許容水準を0.50に設定し、●と△の許容水準を0.1に設定する。この場合、○がとりわけ非寛容的な集団となる。なお、それぞれの集団は各升目に確率0.25で居住するようにし、いずれも許容水準は0.40、移動機会は30回といった条件で30回の試行によって得られた平均を計算している。

まず、一集団だけが寛容的であるようなケースでの分離度の変化は次のようになった。

全体	移動前	平均分離度=0.46
	移動後	平均分離度=0.65
“○”	移動前	平均分離度=0.45
	移動後	平均分離度=0.69

“●” 移動前 平均分離度=0.46
 移動後 平均分離度=0.70

“△”（寛容的） 移動前 平均分離度=0.42
 移動後 平均分離度=0.52

次に、一集団だけが非寛容的であるようなケースでの分離度の変化は次のようになった。

全体 移動前 平均分離度=0.46
 移動後 平均分離度=0.51

“○”（非寛容的） 移動前 平均分離度=0.45
 移動後 平均分離度=0.55

“●” 移動前 平均分離度=0.44
 移動後 平均分離度=0.47

“△” 移動前 平均分離度=0.44
 移動後 平均分離度=0.48

この結果から、両ケースでは移動によって分離・集住が発生する程度が異なってくることをわかる ($t=7.14$, $df=29$, $p<0.01$)。寛容的な集団が少数派になるケースよりも、非寛容的な集団が少数派になるケースの方が分離・集住の発生する程度が低く抑えられている。たとえ、寛容的な集団であっても多数の非寛容的な集団と相互行為する場合には分離・集住のプロセスに巻き込まれてしまわざるをえない。しかし逆に、寛容的な集団数の方が多くなった場合には、仮に全体社会に非寛容的な集団が含まれていたとしても、分離・集住のプロセスを抑え込むことが可能になる。これもまた、＜非＞寛容の限界だといってよいだろう。

4. 議論

最後に、本論の分析結果をまとめ、その含意について簡単に検討することにしよう。

シェリングが明らかにした重要な知見である「比較的寛容な許容水準をもっていても、個体が移動を繰り返すことで明確な分離・集住が発生する」という現象を本論のモデルによっても確認することができた。しかしそれと同時に、許容水準を厳しくすればそれに応じて分離・集住の程度が激しくなるという線形の関係にはなっていないことも確認された。少なくとも本論で用いたモデルでは、許容水準がある段階を超えるとむしろ分離・集住の程度が弱まっていく。いわばこれは、＜非＞寛容の限界と呼んでもよい現象であろう。

また、集団の大きさに差がある場合にもやはり分離・集住が起きる。しかし、分離・集住が形成される過程は、二集団の場合と三集団の場合とでは微妙に異なっている。二集団の場

合には、マジョリティ集団の分離度はあまり変化せず、マイノリティ集団の分離度が大きく変化すること、またマイノリティ集団の分離度がマジョリティ集団の分離度を上回っていることから、マイノリティ集団が集住をはじめることによって分離が形成されていると判断できる。しかし、三集団の場合には、マイノリティ集団の分離度はマジョリティ集団の分離度を下回っており、むしろマジョリティがマイノリティを排除することで分離・集住が形成されていると判断できる。集団数は分離集・集住の過程に影響力をもっており、二集団であればまる知見を三集団以上の場合にもあてはめることは必ずしもできないことが明らかにされた。

激しい分離が生じるのは集団の数が二つだからであり、集団の数が増えることによって、すなわち全体の多様性が増すことによって、分離・集住の程度が弱まるのではないかと期待されたが、少なくとも本論のモデルでは結果は逆になった。むしろ、三集団の場合の方が、激しい分離・集住が発生している。この事実は、全体の多様性が差別的秩序の形成に及ぼす影響を考察する際には、重要な示唆を含んでいると考えられる。

最後に、集団間で許容水準が異なる場合について、すなわち他集団への寛容の度合いが集団ごとに異なっている場合についても検討した。その結果、敵対的な集団に囲まれたとき、自集団の寛容さは分離・集住の程度を和らげる効果をもつが、分離・集住を防ぐことまではできないことを確認した。しかし、寛容な集団が全体の多数を占めるときには、例え非寛容な集団が混じっていたとしても、分離・集住の程度は弱いものにとどまり、ここにも非寛容の限界を見出すことができた。

【注】

- (1) なお、本論でシミュレーションにもちいたモデルは、VBA をもちいて Excel 上でシミュレートされた。
- (2) シェリングのモデルでは、個体は条件を満たす空き地の中で現在居住している升目にもっとも近い升目に移動することを前提にしている。しかし、本論のモデルではそのような前提をおいていない。本論では、最初に見つかった空き地に移動するものとしている。その理由は2つある。ひとつは、本論では Excel の VBA をもちいてシミュレーションを行っているため、計算に負荷がかかりすぎるという難点が生じるためである。もうひとつは、移動する距離に応じてコストがかかると思った場合、許容水準によって移動するか否かを一律に決定するモデルが不適切になるからである。移動のコストを支払ってでも移動したいと考えるための条件が許容水準であるから、移動距離によってコストがかかるとうると移動距離によって許容水準も変化すると考えることの方が自然になる。しかし、この仮定はシェリングのモデルと明らかに異なっているし、またこの条件を反映するようにモデルを変更してしまうとモデルのもっていた単純さという魅力が失われてしまう。

【文献】

- 籠谷和弘. 2004. 「なぜ差別しなくても外国人居住区ができるのか」土場学他編『社会をくモデル>でみる 数理社会学への招待』勁草書房.
- Schelling, T. C. 1978. *Micromotives and Macrobehavior*, W. W. Norton.
- Zhang, Junfu. 2004. "A Dynamic Model of Residential Segregation," *The Journal of Mathematical Sociology* 28(3): 147-170.

最終提案ゲームにおける適応的学習と懲罰

— Abbink 実験の日本における追試の紹介 —

石原英樹

(日本女子体育大学)

【要旨】

最終提案ゲーム(ultimatum game)の実験結果は合理的選択理論の予想(部分ゲーム完全均衡)と大きく異なることが知られている。これに対していくつかの説明がなされるが、どれも決定的ではない。本稿では、学習モデルと、学習ではない懲罰行動モデルを峻別する設定で日本の大学生を対象に実験を行い、先行研究と結果を比較してみた。提案者に関しては、不均等提案を過去の経験で説明する適応的学習モデルを推定した結果、不均等提案拒否の経験はマイナスに有意に寄与し、不均等提案受諾の経験も有意にプラスに効いており係数の値も先行研究を支持していた。加えていくつかの異なる結果がみられたので報告する。

キーワード: 最終提案ゲーム(ultimatum game), 適応的学習(adaptive learning), 公平性(fairness)

1. はじめに

最終提案ゲーム(最後通牒ゲーム the ultimatum game)では、プレイヤー1(提案者)が、ある与えられた総額の分割を相手に提案する。もしもプレイヤー2(応答者)が受け入れたら、総額は第一プレイヤーの提案どおりに分割される。もしもプレイヤー2が拒否をしたら、両者は何も得られない。部分ゲーム完全均衡に至るための遡及的帰納法 Backward induction を用いると、最適反応として、プレイヤー1はほとんど0に近い金額(例えば1円)をプレイヤー2に提案しプレイヤー2はそれを受諾することが容易に導かれる。その結果ほとんどすべてをプレイヤー1が得られることになる。

しかしギュースらの先駆的研究(Güth *et al.* 1982)以来、最終提案ゲームの実験では、部分ゲーム完全均衡予測とかなり異なる結果が出ていることが広く知られている。

例えば提案者の戦略を連続にすると(どのくらいの割合で分割するのかを提案者の意向で変えられるようにすると)提案者の60~80%はパイの0.4から0.5くらいをもちかけ、応答者のうちで0.2以下の提案を受け入れるのはわずか3%であるなどである(Roth 1995)。

この注目すべき実験結果の説明はこの20年間さまざまになされており、主にゲーム理論的説明と分配的公正による説明の対立としてまとめられてきた(Roth 1995)。しかしここ10年間は、議論の対立はやや様相を変え、別の対立軸をめぐってなされるようになってきた。一つはプレイヤー1および2の適応的学習による説明、もう一つは(特に)プレイヤー2の懲罰行動による説明である(1)。

以降行われてきた実験は適応的学習と懲罰の説明のどちらかを決定的にするものはない。

問題なのは、多くの実験がこの二つの説明をうまく分離した実験になっていないということである。議論の整理のためには、適応的学習と懲罰を明確に分離した実験を行う必要があった。

Abbink *et al.* (1996)(2001)の実験はこの分離を行う注目すべき試みである。本報告は、Abbink らの実験を日本で追試したものである。結果の大筋は Abbink らを支持するものであった。しかしいくつかの無視できない相違もみられたので、それを報告する。

2. 先行実験の概要

2.1 懲罰仮説

Abbink, Bolton, Sadrieh, Fang-Fang (1996)(2001)の議論を敷衍してみよう。

最終提案ゲーム実験に関わる理論は基本的に部分ゲーム完全均衡の予測と実際のプレイとの齟齬から出発していることは述べたとおりである。Abbink らは最終提案ゲームにおいて部分ゲーム完全均衡が成立する場合を簡潔に示すため、以下の三つの仮定をおく。

P1: 行為者の主要動機は、最大の金銭的利得を得ること

P2: 行為者は、相手の動機を知っている

P3: 行為者は、自分の最適な行為を計算で求めることができる

すると部分ゲーム完全均衡が実現するという予測は次のように構成される。プレイヤー1は次のように考える。プレイヤー2は「利益が少ないより多いほうを好む」(P1, P2)。最小額をプレイヤー2に与え、残りを自分のものにする提案すべきである(P3)。プレイヤー2は間違いなくこれを受諾する (P1, P3)。

懲罰仮説では、プレイヤー2が、不均等な提案に対する不公平感を和らげるために、拒否をするという説明をする。プレイヤー1は何も得られないことを恐れて完全均衡提案をしない傾向をもつのである。特に Bolton(1991)は選好の個人間比較によってこの仮説を表現している。

2.2 適応的学習仮説

適応的学習仮説ではプレイの動的な性質が最終提案ゲームにおける行動の鍵となる。この理論における重要な研究は、Roth & Erev(1995)によるものと Samuelson(1997)によるものがある。後者については石原・金井(2002)で修正を含めて言及をした。それぞれやや異なった学習アルゴリズムを用いているが、基本的にはほぼ同じモデルである。

ゲームを複数回繰り返すときプレイヤーは最初のゲームでは、どの純粋戦略をプレイするかについての「傾向」を持っていると仮定する。ゲームが繰り返されることで、この傾向は適応過程を通じて変形される。変形率は戦略の過去（前のゲームでの）パフォーマンスに依存している。つまり高い利得を得られる戦略はより高い適応率を持つということである。その結果、プレイヤー2にとっての二つの学習、すなわちプレイヤー1からの高い分配額の提

示の受諾の学習と、低い分配額の提示の受諾の学習では、後者の学習のほうが遅い。前者のほうが高い利得を得られるからである。このことがプレイヤー1を、不平等提案の要求から、プレイヤー2に受け入れられやすい平等提案の提示へと向かわせることになる。すなわち適応的学習仮説が示す、実際のプレイに対する説明は「プレイヤー1が自分の利得が高くなるような戦略をとるよう学習をするほうが、プレイヤー2が不平等提案を受諾する戦略を学習するよりも早い（適応率が高い）」ということになる。

まとめると、適応的学習仮説は、完全均衡仮説のうち、P1を受け入れ、P2とP3を以下のように変更したものだということになる

P2*: プレイヤーはゲームの最初は、いくつかの純粋戦略について、なにがしかの傾向(propensity)を持っている。

P3*: プレイヤーは、P2*の傾向をとる。

完全均衡と比較した場合、懲罰仮説はP2とP3はそのままにP1を以下のように置き換えたものである。

P1*: 相手(プレイヤー1)と比較して非常に小さいシェアしか与えられないとき、プレイヤー2は両者が何も得ないほうを好む。そうでないなら、プレイヤー2は自らの利得が最大になるような分割を選好する。

懲罰仮説は最終提案ゲームにおけるプレイを以下のように説明する。これは標準的な完全均衡の仮説とほぼ同じ方法で構成されている。

プレイヤー1はプレイヤー2の選好を知っている(P1*およびP2)。プレイヤー1は利得が無いよりも、プレイヤー2の許容範囲の提案をすることを選ぶ(P3)。プレイヤー2はこれを受け入れる(P1*およびP3)。

「純粋な」学習と「純粋な」懲罰の背後には、かなり異なった行動メカニズムが存在する。このことが、この2つの明確な分割テストを行うことを可能とするはずであり、これが、Abbink 実験の意図するところである。学習仮説は認知的ダイナミクスを導入する(P2・P3をP2*・P3*に置換)が、厳密な自己利益的な動機(P1)の標準的仮説は保持される。懲罰仮説はこれに対し、プレイヤーの動機に変形を加える(P1をP1*に置換)が、完全均衡仮説と同様な静的なフレームワークで行動を記述することが可能である。

2.3 実験デザイン

Abbink 実験は最終提案ゲームを単純な構造(図1)に縮約した上で、不完全情報を導入する。すなわち二つのゲーム設定(treat)のどちらが行われているのかに関する情報をプレイヤー2だけに伝達するのである。

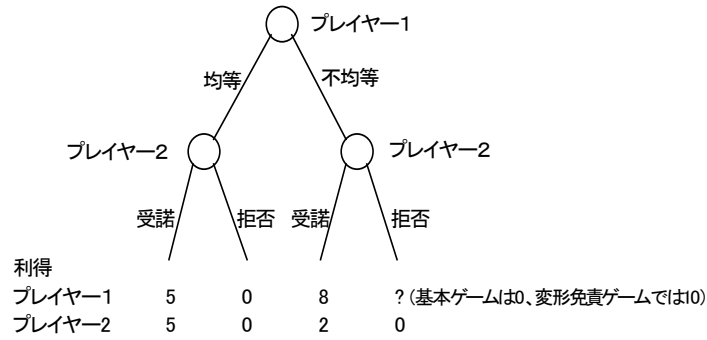


図 1. 報酬に不確実性のある簡易最終提案ゲーム(cardinal ultimatum)

(1)基本ゲーム：先行実験に準拠する．プレイヤー1は「均等な分割 ($a_1 : a_2 = 8 : 2$)」もしくは「不均等な分割 ($a_1 : a_2 = 5 : 5$)」のいずれかを提案することができる．次にプレイヤー2はこれを受諾するか、拒否することができる．

(2)罰なしゲーム (impunity game)：免責ゲーム．プレイヤー1が不均等分割を提案した場合、プレイヤー1は総額 c をすべて受け取る．さらにここに不完全情報を導入する．図2の「?」にあたる部分がそれである．ここはプレイヤー2が不均等な提案を拒絶した場合の、プレイヤー1の利得を示す個所である．

集められた被験者たちには事前に以下のような情報を与える．

(a)?の値はランダムに0か10かのいずれかに決定される：被験者がプレイするゲームは基本ゲームか罰なしゲームのいずれかであり、それはランダムに決定される．

(b)?の値はプレイヤー2にのみ与えられる：すなわちプレイヤー1には現在どちらのゲームをプレイしているかについての情報は与えられない．

(c)実験は何回かのゲームを続けて行い、交渉相手は毎回変わることを．一度交渉したプレイヤーと再びプレイすることはない．

(d)?の値は一回の実験を通じて不変である：被験者は基本ゲームのみを何回か、もしくは罰なしゲームのみを何回かプレイする．

(e)被験者が一度与えられた役割は一回の実験を通じて不変：ある被験者は実験中、プレイヤー1のみ、もしくはプレイヤー2のみを経験すること．

(f)各プレイヤーには実験終了後、ゲームの利得に応じた金額が支払われる．

Abbinkらはこの実験を基本ゲームのみ、罰なしゲームのみでそれぞれ5度(session)ずつ行った．これを以降、Abbink実験と呼ぶ．一回の実験は16人の被験者が参加し、8回(8round)のゲームを1sessionとしてプレイした．各自1sessionのみの参加である．Abbink実験では被験者総数は160人、ドイツの大学の掲示板で募集され大部分が経済学か法学の学部生だった．被験者は指示された場所と時間で1sessionのみの実験に参加した．16人の被験者はランダムにプレイヤー1かプレイヤー2に割り振られ、その情報は各自に伝えられた．実験はPCの画面に図1と同様のツリーの図が表示される．インストラクションおよび画面の詳細はAbbink

et al. (2001)の AppendixA に、全データは AppendixB に掲載されている。?の値はプレイヤー2の画面では 0 あるいは 10 と明示されている。プレイヤー1 が選択するのを受けてプレイヤー2 が画面上で選択をして 1round が終わる。8round でプレイは終了し、結果に応じた金額が被験者に支払われた。

完全均衡の仮説ではこの不完全情報の導入が実質的な意味を持ち得ない。なぜなら?の値はプレイヤー2 の利得には影響を与えないからである。ゆえにプレイヤー2 の行動に影響を与えないはずである。となれば、プレイヤー1 もまた?の値とは無関係に不均等な提案からの利得を期待できる。したがって二つの *treat* の結果に差が出ないはずである。

では適応的学習モデルと懲罰モデルではどうか。

実験の第1セットを考えてみる。このとき?の値に関わらず、プレイヤー1 には同じゲームに見えることから、ここでのプレイヤー1 のプレイには差異は見られないはずである。一方プレイヤー2 は?の値を知っているのだが、適応的学習仮説に従えば?の値はプレイヤー2 の判断に影響を与えないはずである。学習でプレイヤーに影響を与えるのはプレイのパフォーマンスの履歴(自分のペイオフ)だからである。次に第2回は第1回の履歴を参照するのだから、ここでもまた差異は無いはずである。以下同様。適応的学習モデルでもこの二つのゲームの結果に違いはないということになる。しかし完全均衡の仮説とは、両ゲームとも大きく異なっているはずである。

これに対して懲罰モデルは両ゲーム間に差異が出ることを予想する。罰なしゲームではプレイヤー2 が不均等な提案を拒否することは、懲罰として機能しないだけでなく、それを受諾するよりもプレイヤー1 にとっては得になってしまう。となれば、プレイヤー2 は受諾を選択するはずである。従って、基本ゲームと比較して罰なしゲームにおいて、完全均衡がより多く観察できるとの予想が成り立つ。つまり懲罰モデルにおいては相手の情報が重要性をもつ。

3. 再現実験

3.1 実験の概略

Abbink 実験の再現は 2002 年に明治学院大学と早稲田大学で行われた。以後この実験を追試と呼ぶ。

追試では、適応的学習モデルの再検討の意味で *session* を 2 倍にした実験と、懲罰モデルを再検討するために、社会的信頼性や公平性に関する質問紙調査を併用し分析を行った。これらの新しい知見に関しては別稿に展開したので今回は割愛する。

Abbink 実験では *cgi* ではなくアプリケーション (*RatImage software*) を使用しているが、追試では今後の類似実験の便宜を考慮しほぼ同じ内容の *cgi* プログラムを C 言語で開発した。追試に際して契約したレンタル・サーバ上でコンパイルした。レンタル・サーバ上に用意されたコンパイラは *gcc version egcs-2.91.66 19990314/Linux (egcs-1.1.2 release)* である。

追試は全部で 8 回行ない 128 人の被験者を掲示ではなく大学のアルバイト募集の手続きに従って募集した(データはそのうち手続き上のミスがあった 1 回を除いた 7 回を分析に用い

た) . 単位取得などの条件はなく、金銭のみが参加の動機となっている(2).

各 session のはじめに、被験者はインストラクションを読んだ (Abbink 実験を日本に翻訳したもの) . その後、質問紙を配布した (プロジェクトで使用した項目から転載) . 実験担当者が手続きを口頭で説明した. 被験者はくじ引きで決まった番号の机に着席し、各自の持参した携帯電話でゲームを行った. Abbink 実験ではゲームの木が画面に現れるが、追試では簡易画面とした (割愛) . 応答者には?の値をスクリーン上で表示してあるが提案者には表示していない. 1点を20円とし、基本給900円に加えた. Abbink 実験での平均支払い金額は DM18.44 (およそ\$14.75) で所要時間は平均45分. これに対して日本での追試での平均支払い金額は7sessionすべての平均支払い額 (基本給含む) 1588円, 所要時間は平均60分だった. Abbink 実験と異なるのはアプリケーションや支払い方法だけでなく、端末にPCではなく携帯電話を用いたことである(3).

3.2 結果の概要

追試では、通常の最終提案ゲームにおける不均等提案拒否率(15-20%)(Roth 1995)よりもかなり高い拒否率になっている. 以下, Abbink 実験と追試の結果の対比を行う.

?=10	0.0%	1.7%	3.5%	12.0%	12.3%	:平均 5.9%
?=0	3.4%	16.1%	25.0%	25.5%	27.8%	:平均 19.6%
?=10	0.0%		2.8%	11.8%		:平均 5.3%
?=0	43.3%	45.5%	65.0%	72.0%		:平均 54.6%

最初の二行が Abbink 実験の数値(各 session の平均値を小さい順にならべた)である(3). Abbink らの実験では、この5sessionの平均が treatment 間 (?=10 と?=0) で有意差がないとする仮説は棄却される(permutation test 片側 p=0.020). 同様に本研究でも有意差がないとする仮説は棄却された(片側 p=0.000).

図2および図3は、原論文図2に相当する図を、Abbink 実験と追試で比較したものである.

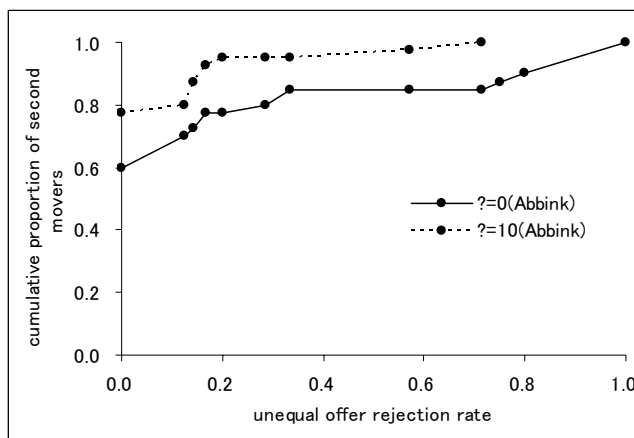


図2 プレイヤー2の不均等提案拒否率の累積度数 (Abbink 実験)

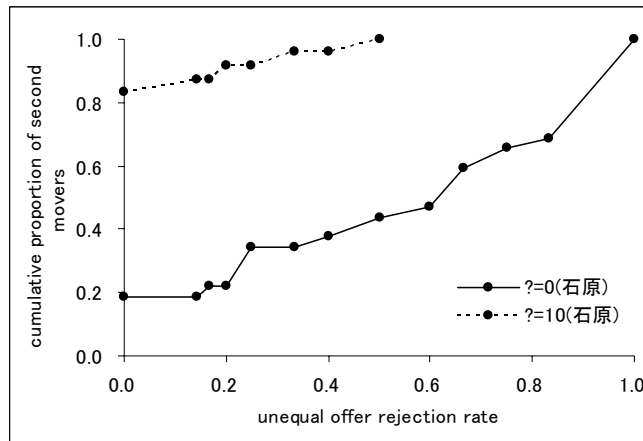


図3 プレイヤー2の不均衡提案拒否率の累積度数（追試）

図2と図3はプレイヤー2の不均衡提案拒否に対する別の側面からの分析である。

図2をみてみよう。プレイヤー2の拒否率の累積度数は $\delta=10$ の場合は $\delta=0$ よりも高い。だが、プレイヤー1からなされた不均衡提案の20%以上を拒否するのは、 $\delta=0$ の場合は24%なのに対して、 $\delta=10$ ではプレイヤー2のわずか5%にすぎない。 $\delta=0$ の場合のプレイヤー2のほうが正の拒否率が高く(z-test 片側 $p=0.046$)、平均拒否率も高い(z-test 片側 $p=0.013$)。 $\delta=10$ で拒否した人9人のうち2人だけが複数回拒絶していた。これに対して $\delta=0$ の場合は16人の拒否のうち9人が複数回拒絶していた。

同様のことを図3（追試）でみると $\delta=0$ の場合の拒否率が大きく散らばっていることが注目される。

もう一つ別の分析をみてみよう。roundごとにみた部分ゲーム完全（プレイヤー1が不均衡提案をし、プレイヤー2がそれを受諾するという組み合わせ）の比率である。

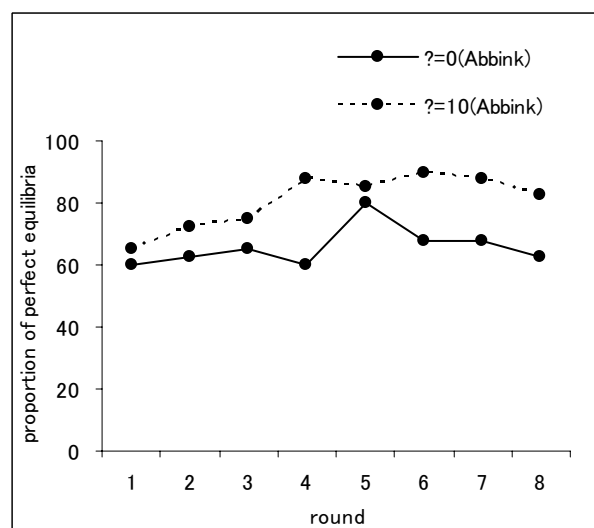


図4 完全均衡プレイの比率(Abbink 実験)

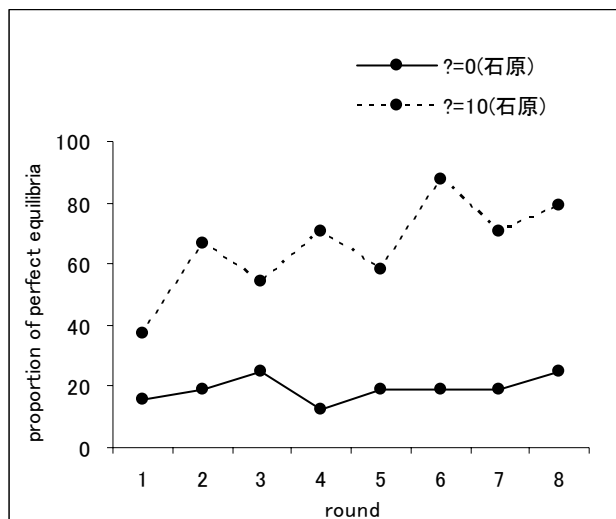


図5 完全均衡プレイの比率(追試)

図4ではすべてのroundで? ϵ =10の場合が? ϵ =0よりも高い。しかしその差はroundの後になるほど目立つようである。Round1ではtreat間の差は明確ではない(0.65と0.60でそれぞれ40サンプル, 片側 $p=0.322$)。しかし最後のround8ではその差ははっきりしてくる(0.83と0.63, 片側 $P=0.023$)。

図5の追試の場合も同様に、すべてのroundで? ϵ =10の場合が? ϵ =0よりも高い。その差は最初のroundから比較的是っきり出ており、後になるとより目立つようになる。Abbink実験と追試の違いはいくつかある。まず第一回目で追試では? ϵ =0と? ϵ =10で完全均衡プレイの比率がかなり異なっていることである。この図からは、懲罰モデルが支持されるだけでなく、? ϵ =10のtreatにおける学習がかなり明確にみられるようである。

次に図6と図7はプレイヤー1と2の行動を分離したものである。

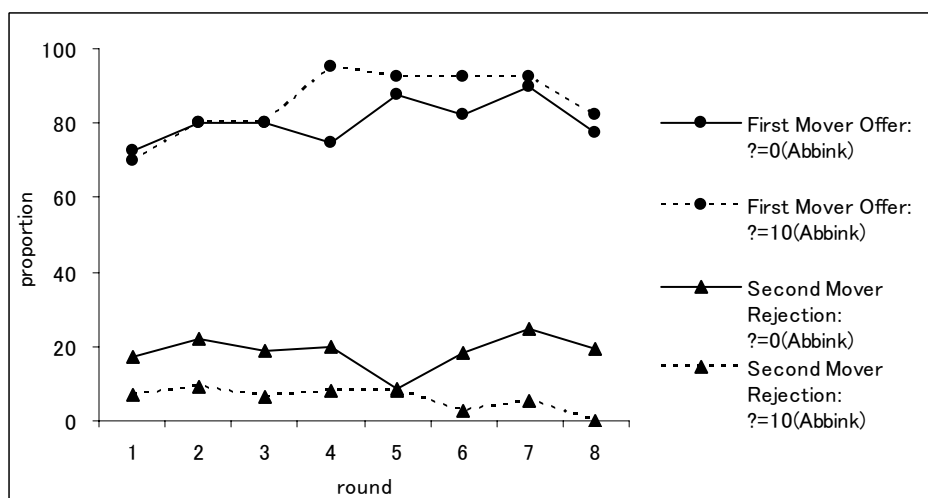


図6 プレイヤー1の不均衡提案とプレイヤー2の拒否(Abbink実験)

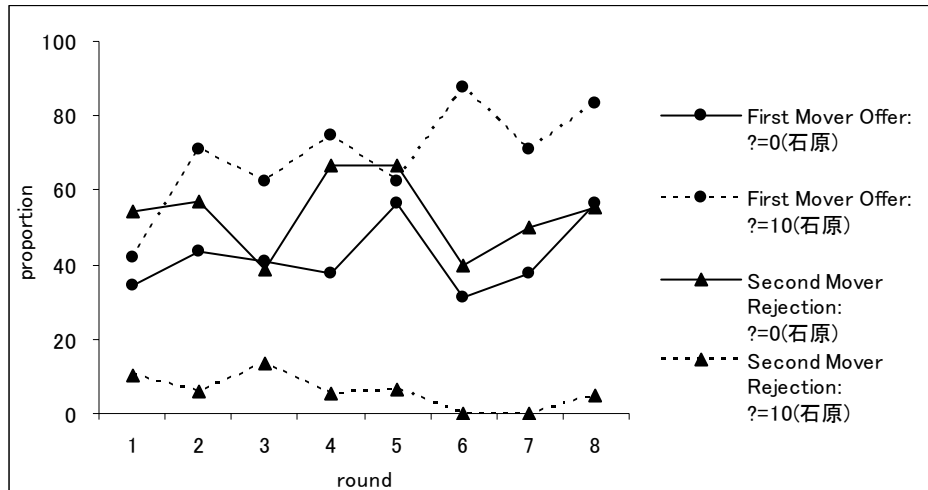


図7 プレイヤー1の不均衡提案とプレイヤー2の拒否(追試)

図6では、 $\gamma=0$ での拒否には round を重ねても trend はみられない。一方 $\gamma=10$ の場合は回を重ねるごとに拒否傾向が弱まってゆくように見えるが統計的に有意ではない(Abbink *et al.* [2001:11])。

ほぼ一貫して $\gamma=10$ のときのほうが $\gamma=0$ のときよりも高いがその差は小さく (86%と 81%) 統計的に有意ではない(Abbink *et al.* [2001:11])。

これに対して図7の追試でも $\gamma=0$ での拒否には変動は大きいものの、round を重ねても trend はみられない。一方 $\gamma=10$ のときには Abbink 実験と同様に拒否傾向が弱まってゆくように見えるが統計的に優位ではなかった。プレイヤー2には学習がみられないということがこの図でもいえそうである。

$\gamma=0$ のときプレイヤー1の trend は Abbink 実験と異なりかなりはっきり上昇しているように見える。総じて追試では Abbink の場合よりも、 $\gamma=0$ の場合と $\gamma=10$ の場合ではプレイヤー1の行動に大きな違いがみられるということである。

ここまでは、 $\gamma=0$ と $\gamma=10$ でのプレイヤー1およびプレイヤー2の行動にいくつかの面で大きな違いがあることを示していた。すなわち適応的学習では最終提案ゲームの実験結果のすべては説明できないということである。

しかし図6や図7では明確ではないにせよなんらかの trend がみられる。そこでプレイヤー1の学習については別の分析からアプローチしてみよう(Abbink *et al.* [2001:12])。被説明変数はプレイヤー1の t 回目の戦略 (不均等提案=1) で、これを $t-1$ 回までのプレイ履歴で推定したのである。履歴は次の4変数である。

$R_{i,t-1}$ = $t-1$ 回目までに不均等提案をして、プレイヤー2から拒否された回数

$A_{i,t-1}$ = $t-1$ 回目までに不均等提案をして、プレイヤー2に受諾された回数

$E_{i,t-1}$ = $t-1$ 回目までに不均等提案をして (受諾された) 回数

$Treat_i$ = $\gamma=0$ ならば 1 $\gamma=10$ ならば 0

Abbink らの先行研究に従い、提案者 i 間の異質性(heterogeneity)を説明するランダム効果を加えたプロビット分析を行い不公平提案を過去の経験で説明する適応的学習モデルを推定した。提案者は合計 56 人いる (Abbink らの実験では 80 人) ので、それらの個人間の差異と、検証したい影響の効果を分離する必要があるからである。

Abbinkデータによるプロビット分析の結果(再現)

Parameter	Estimate	Pr > t
R	-0.49	0.030
A	0.26	0.000
E	-0.04	0.782
Treat	0.05	0.862
Log Likelihood	-218.2	

表 1 Abbink 実験のプロビット分析 (石原による再現)

Abbink 実験に関して、履歴に関する変数の係数の符号は適応的学習モデルの予想通りになっているが、 $E_{i,t-1}$ の係数は有意ではなかった。Treat 間での差異がないことも注意したい。なおこの値は同論文に掲載された原データから再現した値である (論文の掲載値と p 値はそれぞれ、R= -0.51[p 値=0.003], A= +0.26[<0.001], E= -0.01[0.916], T= +0.05[0.862], Log likelihood = -218.2) 。

石原データによるプロビット分析の結果

Parameter	Estimate	Pr > t
R	-0.3975	0.0536
A	0.2592	0.003
E	0.1488	0.1738
Treat	0.5352	0.1335
Log Likelihood	-250.30	

表 2 追試のプロビット分析

追試でも同様に、不均等提案拒否の経験 $R_{i,t-1}$ はマイナスに有意に寄与し、不公平提案受諾の経験 $A_{i,t-1}$ は有意にプラスに効いていた。プレイヤー1の行動には適応的学習がみられるということである。ただし先行研究では公平提案 E は有意ではないが符号がマイナスであったが当該実験ではプラスに効いていた (有意ではない)。プレイヤー2から拒否された経験はプレイヤー1の不均等提案を減少させ、プレイヤー2が不均等提案を受諾した経験はプレイヤー1の不均等提案を増加させている。

先の分析も併せると提案者の行動に関して先行論文との大きな違いは三点である⁽⁵⁾。

- (1) 提案者の均等提案が多い。
- (2) 応答者の拒否率が高い。
- (3) $Treat_t$ が効いている可能性がある (ただし統計的に有意ではない $p>0.1$) 。

(1)と(2)に関しては安易に Abbink 実験との比較はできない。支払い方法や実験端末などで

いくつかの相違がみられるからである。

(3)については今回のデータ分析には含めなかったが、予備的な実験では *treat* の効果が統計的に有意になることが観察された。係数の方向から考えるとこれは適応的学習(*adaptive learning*)だけでなく、ゲームに関する信念を基礎にした学習 (*belief-based learning*)がみられたということを示唆している。

3.3 考察

追試では、Abbink 実験の主な結論、すなわち(1) プレイヤー2 は適応的学習では説明できない行動をしている、ただし(2)プレイヤー1 は適応的学習を行っていることは同じように確認ができた。

提案者の均等提案率や応答者の拒否率の相違については、既存研究で指摘されている文化的な差異かどうか、実験デザインが完全に同一ではないので結論を下すことはできない。

プレイヤー2の「懲罰」に関して、当事者の公平性に関する認識や、一般的信頼性などが関係しているのではないかと予想し、追試では一般的信頼尺度ほかの個人属性を質問紙調査によって尋ねている。これら個人属性の変数による推定と、ランダム効果モデルによる推定による考察を合わせたプロビット分析を別に行なった。

プレイヤー1の適応的学習については、追試で示唆された、プレイヤー1の *belief-based* 学習モデルと適応的学習モデルの比較検討が必要である(5)。

本研究の今後の方向としては、プレイヤー2の行動は学習モデルではない「懲罰」の検討、プレイヤー1の行動は学習モデルの精緻化という平行作業ということになるだろう。

このゲームは合理的選択理論批判の橋頭堡として注目すべきゲームだが、説明モデルが百花繚乱で次第にゲームそのものが見えづらくなっている。過去の実験の再現や追試、データの二次的分析の必要がいまこそあると思われる。

【注】

- (3) 有力な研究としては、例えば Henrich (2000)では、ペルーのアマゾンの狩猟民族 Machiguenga Indians に最終提案ゲームをおこなわせ興味深い結果が得られている。Machiguenga では平均の提案はそれまでの実験よりもかなり低く、メディアンはほぼ 20%提案で、15%提案もしばしば受け入れられ拒否率は非常に低かった。こうした比較文化的調査がここ数年増加している (Henrich, Boyd, Bowles, Camerer, Gintis, McElreath & Fehr 2001)。こうした考察では「適応的学習仮説」と「懲罰仮説」の対立はマイナーな問題とあってよい。また、「懲罰仮説」一つとっても、プレイヤー2の拒否のメカニズムが「公平性規範」なのか、「羨望」なのか、「プライド」(塚原 2003)なのか、などいくつかの競合する説明がある。これらの論文のサーヴェイは紙幅の都合で削除した。より根底的な問題提起として竹村・アントニデス (1998)は最終提案ゲームにおけるフレーミング効果の測定から、序数効用理論の基本仮定そのものが成立しない結果の報告をしている。
- (4) Abbink らの実験では数値 1 を 50pfennigs (日本円だと約 20~30 円)に変換して支払われた。これにならって追試での支払いは、数値 1 に対して 20 円を支払った。しかしこれは最低料金 900 円に加えて支払われたものである。大学での募集の際に最低料金を明示せねばならなかったからである。
- (5) 従来のインターフェースを変更せずに実験を行うには、われわれは実験のために自由に利用できる端末を必要数 (1 回の実験に 16 台) 用意しなければならない。これは大学の情報処理室を使い、レンタルした http サーバに被験者の PC をアクセスさせることを予定していた。しかしセキュリティ・ポリシーなどの適用により利用が制限されることが多く、試行錯誤の結果、インターフェースを携帯電話の web ブラウザ上で利用することとした。現在利用されている携帯電話にはほとんど web ブラウザが

組み込まれており、基本的には web ブラウザが必要とする機能をほとんど満たしている。この変更で必要な作業は CGI プログラムが出力する html をごく一部変更するだけであった。具体的には携帯電話の表示桁数にあった表示に変更し、一般的な携帯電話の web ブラウザが解釈できない html タグ（われわれのプログラムははじめから利用していない）をテストするだけだった。そして、被験者の募集広告に次のような応募資格を付け加えた：「当日、web ブラウザが利用可能な携帯電話を持ってこられる方のみ」。実験者が端末を用意せずとも被験者が個別に用意してくるため、われわれは被験者を募集するだけで実験を実施することが可能になった。被験者に必要経費としてパケット接続料金（試算の上毎回 100 円）を支払ったが、携帯電話の利用は、実験を行う上での制約を取り払い、さらには実験をより機動性あるものとした。結果として、構築した実験システムはインターフェースのコントロールを損なうことなく、コストパフォーマンスと機動性において良好なものとなった。ごく一般的な CGI プログラムが書けるならば、携帯電話上で動作するインターフェースは確実に書ける。この実験システムは制限の多い小規模な実験グループには有効なものである。

- (6) 原論文の誤りを直してある。原論文データは Abbink *et al.* (1996) および Abbink *et al.* (2001) で公表されている。
- (7) プレイヤー 2 の適応的学習は Abbink 実験と同様追試でもみられなかった。追試のデータ、実験手続き、プログラムソースは割愛した。

【文献】

- Abbink, Klaus, Gary E. Bolton, Abdolkarim Sadrieh & Fang-Fang Tang. 1996. "Adaptive Learning versus Punishment in Ultimatum Bargaining". *Discussion Paper no.B-381*. Rheinische Friedrich Wilhelms-Universität Bonn.
- Abbink, Klaus, Gary E. Bolton, Abdolkarim Sadrieh & Fang-Fang Tang. 2001. "Adaptive Learning versus Punishment in Ultimatum Bargaining." *Games and Economic Behavior* 37: 1-25.
- Bolton, Gary. E. 1991. "A Comparative Model of Bargaining: Theory and Evidence." *American Economic Review*. 81:1096-1136.
- Güth, Werner, Rolf Schmittberger & Bernd Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior and Organization* v3: 367-388.
- Henrich, Joseph. 2000. "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga of the Peruvian Amazon." *American Economic Review*. 90(4):973-979.
- Henrich, Joseph, Robert Boyd, Sam Bowles, Colin Camerer, Herbert Gintis, Richard McElreath and Ernst Fehr. 2001. "In Search of Homo Economicus: Experiments in 15 Small-Scale Societies." *American Economic Review*. 91(2):73-79.
- 石原英樹. 2003. 「最終提案ゲームにおける適応的学習と懲罰：携帯電話を用いたゲーム実験」『第 35 回 数理社会学会大会 研究報告要旨集』.
- 石原英樹・金井雅之. 2002. 『進化的意思決定』朝倉書店.
- Roth, Alvin E. 1995. "Bargaining Experiments". in *Handbook of Experimental Economics* (Kagel, John H.; Roth, Alvin E. eds.): 253-348.
- Roth, Alvin E. & Ido Erev. 1995. "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economic Behavior*. 8:164-212.
- Samuelson, Larry. 1997. *Evolutionary Games and Equilibrium Selection*. Cambridge, Mass.: MIT Press.
- 塚原康博. 2003. 『人間行動の経済学：実験および実証分析による経済合理性の検証』日本評論社.
- 竹村和久・ゲリット アントニデス. 1998. 「最終提案交渉ゲームにおけるフレーミング効果の意義：ゲーム理論による記述普遍性はどこまであるのか」『日本グループダイナミクス大会』第 46 回大会発表論文要旨.

【謝辞】

この追実験は、本プロジェクトの実験の手続きに多くを負っています。特に本プロジェクトにおいて開発された設問の使用を快諾していただいたメンバーにお礼を申し上げます。実験のプログラムは笹村良彦氏（早稲田大学）と共同で作成しました。

市民活動の継続要因に関する一考察

－質的調査・分析に基づいて－

中島聡子

(関東学園大学)

【要旨】

近年、市民活動が活発化してきている。市民活動とは自発的で営利を目的にしない市民による活動を指し、ボランティア活動を始め、NPO、NGO、自治会など幅広い活動が含まれる。活動の一般的な定義に「継続性」が含まれるが、法や制度に規定されない自由な活動を継続させる要因については十分な検討がなされていない。そこで本論文では、市民活動に関する質的調査を実施し、活動の継続要因に関する考察を行った。具体的には、活動に関与する主体間の相互作用と継続の関係を調査するため市民活動への参与観察を実施し、(1)援助者と要援助者、(2)援助者と援助者の相互作用を約10ヶ月間に渡り観察し、継続要因に関する仮説を生成した。その結果、(1)活動参加者の行動が内集団びいきである、(2)活動参加者が活動者仲間から内的報酬を得る、という2要因が、活動の継続に影響を及ぼす可能性が示唆された。

キーワード: 市民活動の継続、内的報酬、共感、内集団びいき、質的調査

1. はじめに

近年、市民活動が活発化してきている。市民活動とは自発的で営利を目的にしない市民による活動を指し、ボランティア活動を始め、NPO、NGO、自治会など幅広い活動が含まれる。「市民活動とは、市民による自発的、継続的な社会活動で、営利を目的としないものを指す」⁽¹⁾と言われる様に、活動の一般的な定義には「継続性」が含まれることが多く⁽¹⁾⁽²⁾、これらの活動が一時的なその場限りの現象ではなく、社会に根ざし広く認知される存在になることが求められている。しかしその一方、法や制度に規定されない自由な活動を継続させる要因については未だに十分な検討がなされていない。例えば、ボランティア活動の継続性や組織性を検討した園田は、一定の報酬を得る仕組みの方が、責任のある、継続的、組織的な活動につながる主張がある(園田 1999) ことを紹介し、高木らはボランティア活動を長期継続する人達は活動から恩恵を得ていると感じていることを明らかにした(高木, 西川 2000) が、前者は仮説の域を出ておらず、後者は活動する個人のみに着目した分析であった。そこで本研究では、従来取り扱われてこなかった主体間の相互作用を取り上げる。具体的には、活動に参加するボランティア⁽³⁾が周囲と構築する関係性を調査し、市民活動の継続要因を考察していくこととする。

2. 調査方法

2.1 調査手法

活動に関与する主体間の相互作用と継続の関係を調査するため、参与観察によるフィールドワークを実施した。参与観察法を採用したのは、近年のボランティア研究の手法として、フィールドワークが主流になりつつある（Clary, E. G. & Snyder, M. 1991; 高木, 西川 2000）, アンケート調査は長期的な現象の理解には不向きである（佐藤 2002）などの理由からである。（以下、登場人物の名前はすべて仮名である。）

2.2 フィールドの選択

1. 継続的な活動を行うためには、組織的な活動が中心になること、2. 人の相互作用を研究の対象にしているため、より直接的に人と人が作用しあう活動であることが望ましいこと、などの理由により、平成 17 年度の設定を目指して市民が準備を進める A 県の社会福祉法人の設立現場をフィールドに決定した。法人内容は、精神障害者社会復帰施設で、通所授産施設と地域生活支援センターの併設が予定されている。

2.3 活動の沿革

活動の代表者は B 県在住の中田さんである。血縁者が A 県 A 地区に居住しており、土地の提供を受けられると言うことで、この地区で精神障害者の福祉施設の建設を考案していた。しかし、A 地区では近隣住民の反対に遭い計画の続行を憂慮することになる。その頃、知人を通して、A 県の B 地区で社会福祉法人を運営している小林さんを紹介される。小林さんから、「同じ B 地区で一緒に活動をしましょう」と声をかけられた中田さんは、A 地区から B 地区へと設立予定地を変更し、小林さんの協力の下、法人の設立活動が本格的に開始された。

2.4 活動内容と活動メンバー

フィールドが“法人の設立の現場”のため、主たる活動内容は援助者間で不定期に開かれる集会である。この援助者の集会を便宜上ミーティングと表現するが、内容はインフォーマルである。ミーティングが開かれる場所は、ファミリーレストラン、喫茶店、保健所、活動者が既に運営している福祉施設など様々で、開催時間も特に決まっていない。また、ミーティングの際には司会も議事録もないし、設立施設の内容や運営方法以外の話題も多く語られていた。

フィールドに登場するメンバーは、援助者の中田さんと小林さんを中心にして彼らを取り巻く人々で、固定的ではない。フィールドのメンバーとは、設立活動中に何らかの形で 2 人が関与する人々である。彼等の知人・友人以外にも、設計事務所所員、保健所職員などが登場してくる。また、福祉施設の運営者や職員がメンバー中に存在するため、その施設を利用する人々もフィールドに登場する。本研究では、これらの人々を全てフィールドのメンバーとして考える。

本研究では、2003年2月から10月の約9ヶ月分の活動内容をデータとして用いるが、この間に研究者がフィールドで出会った援助者の数は最終的に12名で、内訳は市民10名、行政職員2名である。市民のうち、活動初期に一度ミーティングに参加しただけの西田さんと、終盤で紹介された前川さん2名を協力者、残り8名を主要メンバーと表現する（図1）。

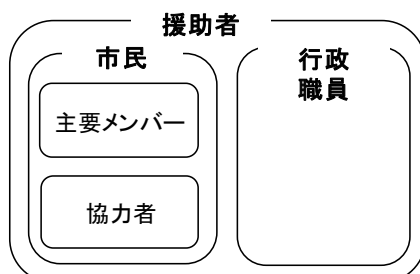


図1 援助者構成図

2.5 調査方法

調査期間は2003年2月から10月の約9ヶ月で、不定期に開催されるミーティングに研究者自身が参加した。時間は長い時ではほぼ1日（8時間程度）、短い時で2時間程度である。初めての参加は2003年の2月で、以後、月に1回～2回のペースで活動に参加した。参与の仕方は、研究者自身も共に活動に参加する active participation⁽⁴⁾ で、活動中は観察データをメモするに留め、帰宅後ワープロソフトを利用しフィールドノートをつけた。基本的には通常の会話の採取が中心で、流れの中でインフォーマル・インタビューを行う時もあった。

3. 調査結果

市民活動に参加して他者を支援するボランティアを援助者、援助者からの支援を受け取る人（何らかの事情により支援を必要とする人）を要援助者と定義し、（1）援助者と要援助者、（2）援助者と援助者の相互作用を観察した。援助者と要援助者の相互作用においては、援助者側からの観察と要援助者側からの観察が考えられるが、本研究では、援助者側からの観察のみ実施した。

3.1 援助者と要援助者の相互作用

フィールドでは、援助者と要援助者の間に、共感と内的報酬の存在が観察された。共感はい援助者の背景から、また内的報酬の生成は、援助者間の会話や援助者と要援助者の直接的やり取りの中から示唆されたものである。

共感の影響

本活動に参加した主要メンバーの背景は、次の通りであった。

主要メンバーの背景(名前:職業, 現在までの精神障害者の人との関わり)

小林:精神・知的・身体障害者のための厚生施設を運営. 精神障害の身内を抱える.

中田:公務員福祉職. 精神障害の身内を抱える.

山田:精神障害者授産施設の職員. 精神障害を抱える友人がいた.

堀田:小林の施設の職員. 精神障害の身内を抱える.

長井:農家. 精神障害の身内を抱える.

坂井:知的障害者厚生施設を運営.

八木夫婦:設計事務所を経営. 過去に, 障害者の福祉施設を設計している.

協力者を除く8名

以上のように, 私的に精神障害者と関係を持つ人が活動者の中に多く存在し(約63%), これらの背景が, 要援助者への共感につながり援助者の行動に影響を及ぼしている事が窺えた. このように, 自分の過去の経験が影響して, 他者への援助が引き起こされる反応は, デイヴィスの組織的モデル(デイヴィス 1999)の中の「役割取得(過程)→応答的共感(個人内的な結果)→援助(対人的な結果)」というルートで説明されている(図2).

デイヴィスの組織的モデルは, 共感概念を包括的に定義するモデルである. このモデルでは, 共感は, 先行条件・過程・個人内的な結果・対人的な結果, という4つのサブグループに分けられる. 「先行条件」とは, 見る側や相手の特質, 置かれた状況. 「過程」とは, 見る側に共感的な結果をもたらすような特定のメカニズム. そして「個人内的な結果」とは, 見る側の中に生じる反応, 「対人的な結果」とは, 相手に対して向けられる行動的反応を意味している. 組織的モデルは, 「過程」や「個人内的な結果」という個人の中の反応が, 援助や攻撃といった相手への対人反応を決定するのに主要な役割を演じていると考えるモデルである.

4つのサブグループには, それぞれ詳細なケースが包含される. 例えば, 「個人内的な結果」には, 並行的, 又は応答的な感情的結果と, 対人的正確さや帰属的判断の非感情的結果がある.

フィールドで観察されたのは, 「過程」グループの役割取得, 「個人内的な結果」グループの応答的共感, 「対人的な結果」グループの援助であった(図2).

役割取得とは, 人が自分の過去の経験と相手の現在の状況を結び付ける能力で, 相手に影響を与えている環境にもし自分が出会ったら自分はどう感じるか, と意識的に想像する力(デイヴィス 1999)のことである. 応答的共感とは「同情や共感, 共感的配慮などさまざまに呼ばれてきたものである. また, 誰かが間違っただけを受けているのを見た際に経験される共感的な怒りもそれである. このどの場合でも, 見る側が持つ感情は相手のそれとは別のものであるが, 相手の経験についての直接的な反応である.」(デイヴィス 1999)と説明される感情的反応である. 最後の援助は言葉の通りで, 他の人々に向けられた利他的な行動の事である.

援助者の一人は, 自分の活動動機を, 次のように述べている.

本活動の動機付け

『もし、自分が同じ立場（精神障害者の人の立場）だったら、と思えばほっておけない。何年間も福祉の仕事をしてきたけど、様々な障害の中でも、精神障害者に関する法制度は一番遅れていて、行政からの援助だけでは障害者の人達を十分ケアできていない、という現状があるの。だから、「精神」の施設を設立しようと思っている。』（中田）

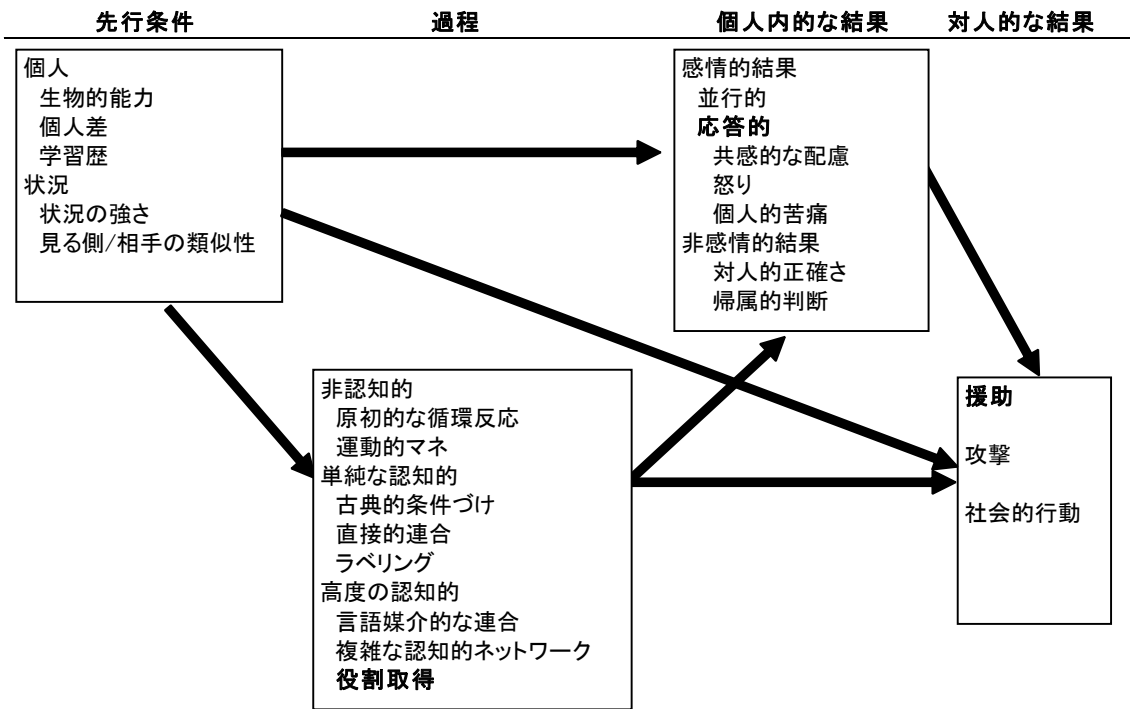


図2 組織的モデル (マーク・H・デイヴィス[1999]より)

「もし、自分が同じ立場（精神障害者の人の立場）だったら、と思えばほっておけない。」という言葉から、役割取得の実行による応答的共感の反応が窺える。また実際に、「精神の施設を設立する」という具体的な援助行動を起こしていることも併せ、先述の「役割取得（過程）→応答的な共感（個人内的な結果）→援助（対人的な結果）」という経路が導き出される。

過去に私的に精神障害者と接する機会があった人々は、同様の障害を抱える人に対して、役割取得や応答的共感が生じやすく、それが援助の原動力になっている可能性があるのではないだろうか。

内的報酬

フィールドでは、援助者と要援助者の相互作用を通し、援助者側に内的報酬が生成される様子が確認された。この内的報酬の存在を記述するため、援助者が要援助者について語った言葉を紹介する。ここでの要援助者は、具体的な誰かではない。各自の中で、より一

般化され抽象化された要援助者に対して向けられた言葉である。

1. 坂井

『私はね、誰かのためって言うんじゃなく、自分の行く道を作っている気がするの。私が助けられていると思う。(福祉を)特別な人がやっているから、みたいな感覚を何とか崩して行きたいと思っている。』

2. 中田

『みんな(要援助者)から、自分も多くの事を学んでいます』

3. 藤井

『一般のボランティアの方たちが援助に入ってくることで、施設外で患者さんとボランティアの人たちが出会ったときに、挨拶や会話が交わされるといった交流が行われるようになったんです。このような交流を通して、両者に相互作用が生じ、双方の心が豊かになっているように感じます。』

4. 前川

『(障害者援助を行う自分について)お金にならないことばかりしてますよー。あははは。』

援助者の「助けられている」「学んでいる」「心が豊かになる」等の発言から、彼らが要援助者との相互作用から、何らかの内的報酬を得ている事が窺える。また、4.の「お金にならないこと」という発言からも、お金ではない何かの存在を感じる事ができる。

3.2 援助者と援助者の相互作用

次に、援助者と援助者のつながりについて、知見をまとめる。活動への参与を通じ、援助者同士が共感した場合、援助者間においても内的報酬が生成される事が観察された。

共感の影響

組織的モデルの「個人内的な結果」に、並行的な共感と命名される共感がある(図2)。並行的な共感は「見る側の者が相手の感情とマッチした感情状態を経験したり、同じ感情がこちら側に生じたりするような感情的状態を経験したりする際に起きる」(デイヴィス1999)ものであり、「見る側とその相手との感情がぴったりとマッチしている」「相手と同じ感情を、それを見る側が経験するようになる」(デイヴィス1999)状態と説明されている。すなわち、並行的な共感とは「自己と他者の感情が一致した状態」と定義できる。

フィールドでは、援助者と援助者の間に、この並行的共感が生じている様子が観察された。

1. 中田と八木

『私も、そのように(中田さんのように)やっていきたい(生きていきたい)と思うけど、なかなか、まだ至らないです。』(八木 妻)

2. 中田と小林

『目標は、小林さんと同じだと感じているわ。(利用者の)「あるがまま」を受け入れることが大事だと感じているの。』(中田)

3. 小林と比留間

『比留間さん(県庁職員)の福祉理念は、わりと近いと思います』(小林)

『近いわね. お仕事という感じじゃない(仕事だから, やっているという印象は受けない)』(中田)

1. の八木さんの発言からは、彼女が中田さんの思想や行動を理解し、中田さんに同調している事がわかる。また、2. の発言からは、中田さんが要援助者に対して抱く気持ちと小林さんが要援助者に対して抱く気持ちが同じものである、と中田さんが考えている事がわかる。これらのことから、八木さんと中田さん、中田さんと小林さんの間で、感情のマッチングが起きている事が確認できる。さらに、3. では、行政職員と市民の間でも、同様の状態が維持されており、このような感情状態は、市民・行政という立場や建前を越え、一個人同士として成立するものである事が理解できる。

この他、自分たち以外の他者を会話の中に登場させる方法で、自分たちの感情の一致を確認する場面もあった。

4. 中田と小林

『私はね, 秋田さん(近隣にある他施設の運営者)とはおそらく福祉活動に対する基本的な考え方が私達と異なっていると思う. だから, あそこはそっとしておく(秋田さんの施設と協力関係は結ばない)ほうが良いと思ってるの.』(中田)

『そう思います』(小林)

秋田さんという人物に対する否定的な意見を述べる際に、「私達と異なる」という表現が用いられている。これは、自分と小林さんの感情の一致を前提にした発言である。この会話から、中田さんと秋田さん、小林さんと秋田さんの間では感情の一致は見られないが、中田さんと小林さんの間では感情の一致が起きている事が窺える。

内的報酬

フィールドではさらに、並行的な共感が生じた各援助者間に内的報酬が生成される様子を確認する事ができた。

並行的な共感がある援助者

1. 八木夫婦と中田

ミーティング終了後、『綺麗な自然を見て行きませんか』(八木 妻)と誘われ、中田さんは八木夫婦と共に近所のドライブへ出かけた。さらに、ドライブの最中に八木夫婦の自宅に招待され、そのまま八木夫婦の自宅に立ち寄った。(研究者もドライブ・自宅訪問に同行)

八木夫婦は、自宅に広めの和室(窓からは、広大な自然が望める作りになっている)を作っており、『誰でも、気軽に立ち寄ってもらいたくて作った』(八木夫婦)と説明があった。『広い事務所にして、事業を拡大しようか』(八木夫婦)とも話し合ったが、『それよりも、お友達に気軽に立ち寄ってもらえる場所にしたい』(八木夫婦)という気持ちで、このような部屋を作ったと話していた。ま

た、この日は『うちにも遊びに来て下さい。主人ともぜひあって欲しいです。』（中田）『よろこんで。』（八木）と言う会話が交わされていた。

2. 中田と小林(1)

ミーティング終了後、中田さんに対して小林さんが『なんだか、楽しくなってきました』と、静かな笑顔でつぶやく姿が観察された。

3. 中田と小林(2)

『中田さんと坂井さんは雰囲気似ている』『会わせたい』（小林）

と言われ、坂井さんが中田さんに紹介された。坂井さんは同地区で知的障害者の厚生施設を運営している。坂井さんのお宅を訪問し(研究者も同行)、昼食と歓談の時間を持ち、坂井さんの施設を訪問した。

昼食時の会話の一部

『(自分の過去の経験では)利用者同士の関係よりも、親御さん同士の方が大変だったかな・・・と思いますね。例えば、うちの子の方が障害の程度が軽い、とか。』（中田）

『(教師時代に)生徒を序列化しなければいけないのを、どうしても悩んだね。』（坂井）

『おれも、作業所やってるとき、俺の人生終わったかな、と思ったよね。世の中が変わらなければ、未来が見えないと思った。』（小林）

『小林さんの頃なんて、高度経済成長期の真っ只中だったんじゃない？』（坂井）

『そうなの。高校 400 人いて、福祉の大学に進んだのなんてたった 4 人だったからね。勝ち組、負け組みって言ったけど、俺なんか本当に負け組みって言う感じだったの。』（小林）

『(序列の問題について)いろんな事が同じ土俵に乗るようにしないとイケないのかな。(これは)社会福祉だけの問題じゃないよね』（小林）

『ないよね!』（全員）

このように、並行的な共感がある援助者達は、相互に好意を抱き、プライベートで遊びに行く、知人を紹介するなどの肯定的な行動を起こし、そこに、友情や連帯感のような内的報酬が存在している事が窺えた。例えば、1. の八木さんは、「お友達に気軽に立ち寄ってもらえる場所にしたい」と説明した自宅の一室へ中田さんを招待している。これは、中田さんが八木夫婦から友達と認知された事実で、両者の間に友情という報酬が生成された事を示唆している。また、2. の「楽しくなってきた」という発言は、要援助者を援助する行為についてではなく、仲間ができた事に対する感想である。このことから、要援助者との相互作用が生じない時にも、援助者同士が相互作用する事によって、何らかの内的報酬が生成される事が窺えた。

次に、並行的な共感が生じなかった援助者間の関係を紹介する。

並行的な共感が無い援助者

『秋田さんは、「A 県で一番初めに精神障害者のための施設を建てた、先駆者だ」って(自分のことを)言っていたわ。同じような内容で、近い場所で活動しているんだから、みんなで一緒にがんばれば良いと(私たちは)思っていて、山田さんも挨拶に行ったの。そうしたら、秋田さんに「私の施設で働かないか？」(引き抜きの相談)って言われちゃって、山田さんは帰ってきて、「こっちは、みんなで一緒にいい活動を、と思っているのに、何なんだよ、あの人は？」って、怒っていたわ。』
(中田)

この発言から、山田さんと秋田さんの間には並行的な共感が生じなかった事がわかる。また、その話を聞いた中田さんと秋田さんの間にも、並行的な共感が生じていない。そして、結果的に、かなり否定的な感情が生まれてしまっている事がわかる。このように、並行的な共感がある援助者間には内的報酬が生成されるが、並行的な共感が無い援助者間には、相手に対するネガティブな意見やイメージが生まれてしまう事実が観察された。

3.3 相互作用モデルの構築

前節では、援助者と要援助者、援助者と援助者の相互作用を説明した。援助者と要援助者の間には応答的共感が、援助者と援助者の間には並行的共感が生じ、それぞれの共感した相手との相互作用からは内的報酬を獲得していることが示された。

次に、援助にかかるコストについて考察する。実際に援助を実行する場合、援助者側は施設建設費や運営費と言った援助コスト(外的資源)を負担する事になる。本フィールドでは、授産施設で生産する農産物の販売や補助金制度の利用によりそれらを捻出する予定である。このように、援助者と要援助者、援助者と援助者の間には、内的報酬の獲得以外に、外的資源の移動も生じる。そこで、相互作用によって増減する内的報酬と外的資源の関係を、相互作用モデルとして図3に表した。

相互作用モデルは、援助者と援助者、援助者と要援助者という2つの関係を同時に示すモデルである。並行的共感が生じた援助者同士は、内的報酬を獲得しながら(図3:援助者間の白矢印)援助のための外的資源を生産していく(図3:援助者間の黒矢印)。また、応答的共感が生じた援助者と要援助者の間では援助が実行される。その際、援助者から要援助者へは外的資源が提供され(図3:要援助者への黒矢印)、援助者は内的報酬を獲得する(図3:要援助者からの白矢印)。

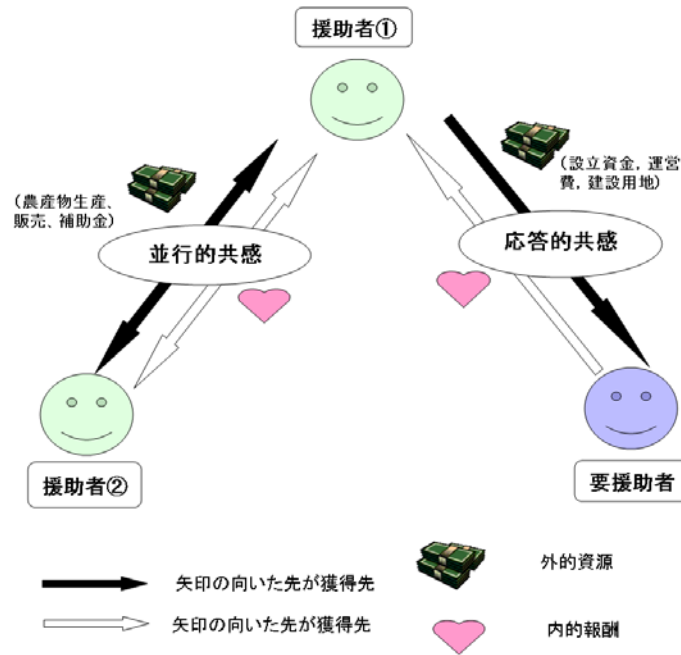


図3 相互作用モデル

4. 市民活動の継続要因—仮説の生成—

以上の観察結果を基に、市民活動の継続に影響を与えるものとして、援助者が要援助者に感じる応答的共感、援助者と援助者が生成する内的報酬という2要因を抽出し仮説を生成した。それぞれの要因が、どのように継続に影響するか、観察結果を含めて説明していく。

4.1 応答的共感による内集団びいき

従来、共感是他者への援助を生起させると言われ、様々な研究が行われてきた⁽⁵⁾。しかし研究者はフィールドで、共感に対する異なる知見を得ることができた。それは、応答的共感の有無が、内的報酬と援助コストのバランスに影響を与える可能性である。

フィールドの援助者達は、応答的共感を覚えた相手(精神障害者)に援助の実行を予定しているが、現実社会には、この他にも他者の支援を求める人々は数多く存在している。多種多様な人間が存在する現実社会で、人々の抱える問題や望む支援の形は様々だ。これらの中から、特に応答的共感を覚える相手を選択して援助を実行する事は、結果的に彼らから援助を受けられる人と、受けられない人という2種類の集団を作り出す事になるが、このような集団をサムナーの言う内集団と外集団と考えることができる。内集団とは、ある個人がそこに所属し帰属感や愛着心を持ち、そこに所属する人々を「われわれ」として意識しうる集団のことである。一方の外集団とは、違和感や敵意を持ち、そこに所属する人々を「かれら」としてしか意識しえない集団のことである(森岡他 2000)。援助を積極的に提供したい相手とそうではない相手として区別すると、援助者達が前者とつくる集団を内集団、それ以外を外集団と考えることができる。そこで、誰にでも利他的に振舞う行動を博愛主義、限られた相手に利他的に振舞う行動を内集団びいきと定義して、2つの行動の効用を考察する。援

助者が博愛主義行動を取る場合、要援助者の選択は行わず全ての人に平等な援助を実行する事になる。その結果、要援助者との相互作用の機会が増え多くの人から内的報酬を得ることになるが、援助コストの負担は増加する。一方、内集団びいき行動を取る場合、共感した相手のみを援助するため、要援助者との相互作用が減少し、援助コストの負担は減るが、その分得られる内的報酬も減少する。このように、博愛主義行動なら多様な人々と相互作用し、多くの喜び（内的報酬）を獲得できるが、コストの負担が大きくなる。内集団びいき行動なら、コストの負担は小さいが、多様な人々と相互作用する機会が減り、喜び（内的報酬）を得る機会が減ってしまう（表1）。

表1 応答的共感による効用

	博愛主義	内集団びいき
内的報酬の獲得	有利	不利
援助のコスト負担	不利	有利

このように2つの行動は異なる効用と欠点を持つので、内集団びいきが本当に市民活動の継続に効果を及ぼすかどうかは自明ではない。しかし、フィールドの援助者たちが、内集団びいきを採用していた事を受け、「応答的共感による内集団びいきが援助コストを低減させ、それが市民活動の継続に良い影響を及ぼす」という仮説を生成することができる。その一方で、本仮説には「獲得できる内的報酬が少ない」という弱点も含まれており、この点についての検討を十分行う必要があるだろう。

4.2 並行的共感を覚えた援助者間で獲得される内的報酬

フィールドの援助者が並行的共感でつながり、両者に内的報酬が生成される様子が確認されているが、彼らの間では、さらに、次のような発言も見られている。

共に活動している事を強調する発言

1.計画開始を決定した会話から

『一緒にやりましょうよ』(山田, 小林)

『(一呼吸して)わかりました。一緒にがんばりましょう』(中田)

2.活動に対する意欲

『小林さんと一緒だから、やろうと言う気持ちになったわ。』(中田)

活動仲間を気遣う発言

【発言の背景】当初は施設を新規建設する予定であったため、八木夫婦が、施設の設計という形で活動に参与していた。しかし活動を進めるうちに、適当な土地が見つからない事、小林さんが小規模事業を押ししている事などが影響し、新規建設から建売物件の購入に話が変わった。建売物件の購入になる場合、設計の必要は無くなり八木夫婦の役割が消滅する。このような状態になったときに中田さんから発せられた言葉である。

3. 仲間意識

『建物を購入するなら(建設しないなら), 八木さん達の事も考えないといけないと思ってるんだけど』(中田)

援助者の1. 2. の発言から, 援助者が活動そのものに対する意義以外に, この仲間と共に活動する事にも意義を感じている事がわかる. また3. からは, 以前の通りの役割分担が無くなったとしても, 相手との関係維持を求めようとする姿勢が読み取れる. これらの事実から, 援助者同士の関係性が, 市民活動の継続に影響を及ぼしていることが窺える. これはブラウが「価値と利害を共有する人びとは, 相互に信念を支持し合い共通の努力にたずさわる基礎をもつ. したがって, 同質的な志向と態度が人びとを互いに引き寄せる. それに対して, 重要な問題における意見の不一致は, お互いを非誘因的な仲間にしてしまう. 」と述べた現象を連想させる.ブラウは, 価値を共有する仲間はその人にとって誘引力を持つと言い, このような仲間に関心がある意見が支持されることはその人にとっての報酬にあたると述べている(ブラウ1974).

さらにここで, ボランティア意識調査の結果を参照してみると, 1位「新しい人間関係ができた」2位「思いやりの気持ちが深まった」となっている(工藤他1998). 1位が援助者間相互作用で獲得される内的報酬, 2位が要援助者との相互作用で獲得される報酬となっており, 援助者間で生成される内的報酬の重要性が理解できる. また, 実際に本フィールドでも, 友情の萌芽・知人の紹介と言った人間関係の構築につながる現象が観察されている.

以上の結果を基にして援助者間の相互作用と活動の継続に関する肯定的な仮説を唱えることができる. それは, 「並行的共感を覚えた援助者同士の相互作用で生成される内的報酬が市民活動の継続に良い影響を及ぼす」という仮説である. すなわちこれは, 要援助者との相互作用で生成される内的報酬だけでは援助者の満足を十分満たす事ができず, 活動の継続は導出されないと言う予測を意味している.

5. おわりに

本稿は調査によって, 仮説(1)「応答的共感による内集団びいきが援助コストを低減させ, それが市民活動の継続に良い影響を及ぼす」, 仮説(2)「援助者同士の相互作用で生成される内的報酬が市民活動の継続に良い影響を及ぼす」を生成し, (1) 応答的共感による内集団びいき(2) 並行的共感を覚えた援助者間で獲得される内的報酬, という2要因が継続に影響を及ぼす可能性を抽出した. また, シミュレーション実験からは, 2つの仮説を支持する結果が得られている⁽⁶⁾.

しかし, 今回の調査では, マクロな要因の影響に関する知見については得ることができていない. 今後は, この点に関する十分な調査が必要である. 是認や承認などの社会的要因が個人の行動に影響を及ぼすことは, 先行研究において指摘されている(ブラウ1974)が, 本調査では, これらの影響を発見するに至らなかった. 調査期間が比較的短かったことや調査実施地でのコミュニティへの関与の度合いなど様々な要因が考えられる. 今後はその地域

特有の文化や規範、習慣などが及ぼす影響を抽出できるよう、微視性と全体性の両方の影響を観察できるような調査設計を行い、それらの社会的要因が個にどのような影響を及ぼすかについても検討を行いたい。

【注】

- (1) 大阪府生活文化部 <http://www.pref.osaka.jp/fukatsu/vngroup/sisin02.html>
- (2) 内閣府国民生活局による市民活動団体等基本調査
<http://www5.cao.go.jp/seikatsu/2001/0409shiminkatsudou/main.html>
- (3) 本研究では、人の関係性に着目して、ボランティアを「困難に直面する他者に働きかけ、つながりをつくる人」とする金子の定義（金子 1993）を踏襲し、無償性は問題としない。
- (4) ボランティアとして参加するなど、フィールドで役割を持ちながら観察する参与方法。
- (5) Toi&Batson は、実験操作によって共感を高めるグループと高めないグループを作って援助の出現率を比較し、共感を高めたグループの方が援助率が高まる結果を得ている（Toi, M. & Batson, C. D. 1982）。Krebs は、他者との類似性が高いと教示して共感性を操作し、それが援助を促進する事を示している（Krebs, D. 1975）。
- (6) 著者は、エージェントベースシミュレーション実験を実施し、本仮説を支持する結果を得ている。詳しくは、中島（2005）を参照。

【文献】

- 園田恭一. 1999. 「地域福祉とコミュニティ」有信堂, pp. 98-99
- 高木修（監）・西川正之（編）. 2000. 「シリーズ 21 世紀の社会心理学 4 援助とサポートの社会心理学」北大路書房, p. 89; pp. 3-6
- Clary, E. G. & Snyder, M. 1991. "A functional analysis of altruism and prosocial behavior: The case of volunteerism." *Review of Personality and Social Psychology*, Vol. 12, 119-148, p. 122
- 佐藤郁也. 2002. 「組織と経営について知るための実践フィールドワーク入門」有斐閣, p. 157
- マーク・H・デイヴィス（著）菊池章夫（訳）. 1999. 「共感の社会心理学」川島書店, p. 14; pp. 46-47; pp. 21-22; p. 124; p. 21
- 森岡清美・塩原勉・本間康平. 2000. 「新社会学辞典」有斐閣
- ピーター・M・ブラウ（著）・間場寿一・居安正・塩原勉（訳）1974. 「交換と権力 社会過程の弁証法社会学」新曜社, pp. 60-62; pp. 229-253
- 工藤敬吉・杉本正治「ボランティア像」大災害で変貌—“気軽型”から“献身型”へ—, 放送研究と調査, NHK 放送文化研究所, 1998年3月号, pp. 26-39
- 金子郁容. 1993. 「ボランティアもうひとつの情報社会」岩波新書 p. 65
- Toi, M. & Batson, C. D., 1982. "More evidence that empathy is a source of altruistic motivation." *Journal of Personality and Social Psychology*, 43, 281-292
- Krebs, D., 1975. "Empathy and altruism." *Journal of Personality and Social Psychology*, 32, 1134-1146
- 中島聡子・中井豊・古宮誠一. 2005. 「エージェントベースシミュレーションを利用した市民活動継続要因の考察」, 情報処理学会論文誌「知の共有から知の協創へ」特集号, vol. 46 No. 1
- 箕浦康子（編）. 2001. 「フィールドワークの技法と実際」ミネルヴァ書房
- 一番ヶ瀬康子・小林博・馬場清（編）. 2002. 「実践・福祉文化シリーズ第4巻 地域社会と福祉文化」明石書店

【謝辞】

本研究はフィールドの皆様方のご協力がなければ成り立たないものでした。心より感謝致します。また、本研究を進めるにあたって多大なアドバイスを頂きました、芝浦工業大学大学院の古宮誠一先生と芝浦工業大学システム工学部中井豊先生、進化ゲーム理論研究会の帝京大学経済学部大浦宏邦助教授と皆様方に感謝の意を表します。

ハビトウス変異とシグナリングのダーウィニアン社会学・序説

一方言変異・変な流行・バビトウス論懷疑から、パスワード改訂仮説・パラサイトシグナリング・多重シグナリングへの暗黙のご了解モデルへー

桜井 芳生
(鹿児島大学)

【要旨】

筆者は、人間の属性について、ダーウィン生物学をはじめとする自然科学的認識を直接援用しようとする社会学を試行する。いままで、人文・社会諸科学において、相互にほとんど関係がないとおもわれていた諸現象について、このアプローチは、非常に蓋然性の強い関連をしめしてくれることが多い。本稿では、とくに、ダンバーとハリスの理説を援用する。方言（とくに、方言の変異）、若者などに見られる「変な流行」、ブルデューが論じた「ハビトウスによる暗黙的選抜」論、さらには、情報の経済学で論じられる「シグナリング」、などが、この視点をとることで、相互に非常につよい通底性をもつことをみとることができる（「パスワード改訂」仮説）。さらには、この視点を「応用」して、「パラサイト・シグナリング」モデルといった社会をみるうえでの新たな視点の提案を試みる。

キーワード: ハビトウス、シグナリング、言語変異、方言周圏現象、パスワード改訂現象。

1. はじめに

現代ダーウィニアン生物学が、社会学さらには広く人文・社会科学全般にもたらしうるインパクトを、進展させる作業を、筆者は最近おこなっている。社会学をはじめとする人文・社会科学は、過去一世紀ほどにおいて、自然科学における「モデル（考え方）」を自己の領域に応用することについては、ある程度熱心であったようだ。しかし、そもそもヒトはどのような属性をもったエージェントなのか、という点においては、少数の例外を除いては、自然科学の成果に直接依拠するアプローチは少なかったと感じられる。

過去一世紀の人文・社会科学の多くがこのような傾向をもっていたこと自体、ダーウィニズム生物学をはじめとする自然科学的人間理解によって、説明することが近々可能になると、私は予感する。自然科学の圧倒的な成功にたいする劣等意識、他方において、そうでありながらも、人文・社会科学を自然科学には還元したくない・人間を他の生物と同等に扱いたくない、といったいわゆるバイオフィビア（生物学的視点の嫌悪）、おそらくこの両者がいまわって、このような傾向が生じた、と感じられる。筆者自身、社会学者をまえにして、自分の「ダーウィン生物学の成果に依拠した、ダーウィニアン社会学」の構想をかたると、ほとんどつねに、「モデル（考え方）」の部分において進化論的発想を利用した進化ゲーム論のようなアプローチですね、と「誤解」されてしまう。

筆者がめざす「ダーウィニアン社会学」とは、そうではなく、人間の属性について、ダー

ウィン生物学をはじめとする自然科学的認識を直接援用しようとする、社会学である。

このようなアプローチを指向する理由・利点はいくつかある。ここではとくに、以下の一つの理由を強調してみたい。

すなわち、いままで、人文・社会諸科学において、暗黙にあるいは明示的にきづかれていた諸現象のうちで、相互にほとんど関係がないとおもわれていた諸現象について、このアプローチは、非常に蓋然性の強い関連をしめしてくれることが多い、という理由である。

本稿では、そのうち、方言（とくに、方言の変異）、若者などに見られる「変な流行」、ブルデューが論じた「ハビトゥスによる暗黙的選抜」論、さらには、情報の経済学で論じられる「シグナリング」、などが、相互に非常につよい通底性をもつことをみてるつもりである（「パスワード改訂」仮説）。

そして、以上のように「すでに見いだされている諸現象」相互に関連をつけるだけではなく、さらには、この視点を「応用」して、「パラサイト・シグナリング」モデルといった社会をみるうえでの新たな視点の提案を試みてみたい。

2. 現代ダーウィニズムの一つとしての、ダンバーの「方言変異」論

本稿で、筆者が依拠したい現代ダーウィニズムとは、第一に、ロビン・ダンバーのいわば「方言変異」論である。方言について、世の人の前でかたると、往々にして「現代においては、さまざまな方言が消滅の危機に瀕していて、、、」とかいった反応が返ってくることが多い。しかし、ある科学的視点からすると、そもそも、方言というものが存在するという事実こそが説明を要する「謎」であるようにみえてくる。なぜなら、言語というものは、その言語を使用する人が多ければ多いほど、その利用者各人にとって便宜が増大するという性質をもっているからだ（ネットワーク外部性）。だとすれば、人類が言語を使用しはじめてこれほど長い歴史がすぎているのに、これほど、多くの言語（方言）が存在しているということとはとても不思議なことにみえてくるだろう。

このような謎にたいして、ロビン・ダンバーが、おもしろいことをいっている。

「方言にはもう一つの長所がある。少なくとも世代という尺度からみれば、比較的短期間のうちに変化することができるのだ。移住した群れは、一世代くらい経つと、同じ言葉を使っても話し方やアクセントを変化させている。オーストラリア移住者のはほとんどが過半二〇〇年の間にそこへ移ったという事実にもかかわらず、オーストラリア人と英国人のアクセントが現在どれほど異なっているか、考えてみよう。そうすると、明らかに提唱できるのは、方言はフリーライダー問題に対処するための適応形態だということだ。群れは新しい話し方、まったく同じことがらに関する新しい言い方を絶えず開発することによって、その一員を確実にたやすく見分けられるようにしている。」（Dunbar1996=1998）

彼ダンバーは、現代ダーウィニストの一人であるので、ハミルトン流の血縁的距離が近いもの同士の相互協力が存在しうることをみとめる。しかし、血縁的近親者による相互扶助には、よそ者がうちわ者を装って入り込んで、協力されることの「利得」のみを得る、という「ただ乗り（フリーライダー）」問題が生じてしまう。このフリーライダー問題への対処戦略

の一つとして、方言変異が進化したのではないかとダンバーは考えるわけである。

3. 現代日本若者における、変な流行、と、受けない化粧

本稿が第二に依拠する現代ダーウィニズムは、ジュデイス・リッチ・ハリスの「グループ・ソーシャリゼーション」理論である(Harris1998=2000)。彼女は、親が子供に文化的な影響を与えるという常識に反対し、子どもの仲間集団による影響が子どもの社会化には大きな力を持っていると主張する。彼女は言う。

「文化を改変できるのは十代後半から二十代前半で独自集団をもつ人々だ。集団性により、親や教師の世代とは一線を画したいという気持ちが駆り立てられる。自分たちは一世代前とは違った存在でありたいという気持ちがあまりに強いため、両者の違いは改善する方向へ向かうとは限らない。実際、改善ではない場合がほとんどだ。彼らは違う行動、違う思想を受け入れ、新語を造語し、斬新なファッションも生み出す。さらにこれらの行動、思想などを携えたまま成人になる。彼らは自分の子どもに差別化を図る新たな方法を生み出すという重圧を残すのだ。ママとパパはマリファナを吸っていたらしい。嫌だ嫌だ、私たちは別のものを吸いましょう！」(Harris1998=2000)。すなわち、子どもや若者においては、仲間集団への準拠が大きく、仲間集団への準拠をおこなうために、その集団の外部と無理にでも差別化をおこなう傾向があるというのである。

4. 仲間集団準拠と「パスワード改訂」仮説

以上の、ハリスの「仲間集団影響説」と、ダンバーの、「よそ者峻別マーカ―」としての「方言変異」説とを、総合して、たとえば若者文化について、以下のように仮説をたてることができるのではないだろうか。

すなわち、若者は同世代の仲間集団の影響を強くうけるのだが、「なかま/そとま」の峻別マーカ―として、「変な流行、や、ファッション」を無意識に「利用」している、と。筆者は、別のところで、昨今の日本の若い女性の「化粧」現象について分析したことがある。とくに「男受けしない化粧」、とりわけ「おじさん受けしない化粧」について、分析したことがある。そこでの主旨は、男うけ（特に、おじさん受け）しない化粧とは、この現象の一例ではないかということであった。

このような現象が、「なかま/そとま」マーカ―としてはたらくためには、「予想もできないような方向に変化」しなければならない。（そうでないと、「外部者」峻別としてつかえない）。

いわば、これらの現象は、「パスワード」の改訂になぞらえることができるのではないだろうか。パスワードはまさに、あるメンバーシップをもつものであるか否かを峻別するフィルタリング機能を担っている。しかし、パスワードは、改訂しないでずっとおなじものを使いつづけていると、やがてはなにかの拍子に第三者へと漏洩してしまう。よって、パスワードはある程度以上の頻度で改訂しなければならない。しかし、この改訂の方向は、第三者にと

って予想のつくような方向では、パスワードの機能を果たさない。

まさに、これと同様に、化粧や変な流行や方言は、第三者が予期しえないような方向に一定程度以上の速度で、変異しつづけねば、フィルタリングの機能をはたさないのだろう（パスワード改訂仮説）。

5. ブルデュー・バビトウス論への一疑問から、方言圏論へ

以上のような視点は、有名なブルデューのハビトウス（習慣）による選抜論への、ある一つの疑問にたいして、かなり答えてくれると思われる。

ピエール・ブルデューの一連の社会学的試みについては、いまではかなり有名となった。彼ブルデューの真意が那邊にあったかは、かなり研究されつつあるといえる。しかし、ブルデューの問題意識とはとりあえず独立に、彼が提起した概念装置・仮説の意義を検討する作業はあまりなされていないように感じられる。

とくに、彼が提起したいわゆるハビトウスによる選抜理論の、意義と問題点（その理論があらたに提起してしまう疑問点ならびにそれへの回答）は、ブルデュー自身の問題意識と独立に、探求されているとはいいがたい、とおもう。

われわれの問題意識にふれあうかぎりにおいて、ブルデューのハビトウスによる選抜論の骨子は以下のようになる。すなわち、

ヒトは、幼児期の社会階層の高低におうじて、階層特有のハビトウス（習慣）を（親などから）無意識に習得する。そして、若年・成年に進んでいく際に、タテマエにおいては学力などで選抜すべき入学試験などにおいて、このバビトウスの差異によって、ヒトは高い階層へのチャンスの多寡へと選抜されてしまう。その結果、全体を客観的にみると、金持ちの子が、金持ちの子であるがゆえに金持ちにふさわしい身のこなし（ハビトウス）をみにつけ、そのおかげで金持ちになりやすい学校などに進学でき、やがては金持ちになった、すなわち、階層的再生産が生じた、だけであった。のに、当事者の視点からは、「勉強したら、あるいは、生まれつき頭よかったから」いい学校に入れて金持ちになれた、と事態が「正当化」されてしまう、わけである（文化資本の無自覚的相続による、社会階層の誤認的再生産）。

このモデルは、非常に興味ぶかい。しかし、素朴な疑問を喚起してしまう。なぜ、貧乏人は、いつまでも、自分の子孫をふたたび貧乏にしてしまうような、貧乏人ハビトウスの相続を継続しているのだろうか？、と。

この疑問は、ブルデューの視点からは容易に解かれてしまうようにみえるかもしれない。すなわち、たしかに、貧乏人が自分の子どもに貧乏人ハビトウスを相続させるは不利な行為だ。が、まさに自覚されず無意識にそうしてしまっているのだ、と。

たしかに、無意識・無自覚性がこのメカニズムにおいて効いていることについては、みとめるのに筆者もやぶさかではない。しかし、無意識性もちだすだけで、この疑問が全く払拭されてしまうとも思えない。なぜなら、ヒトをふくむ動物の行動はほとんどは無意識・無自覚におこなわれている。そうでありながらも、少なくとも平均的長期的には、各個体は自分の不利なるようなことは回避し、自分の有利になるようなことを選択する能力をもってい

るとおもわれるからだ（もし、そのような能力をもっていなかったら、彼の遺伝子は、進化史のなかで、サバイバルしなかっただろう）。

というわけで、上の疑問はいまだ残存することとなる。われわれとしては、すぐのちに「合流」するが、さしあたり「二手に分かれる、二本の流れ」によって、この疑問へ回答を提起してみたい。すなわち、第一が上述の「パスワード改訂現象」であり、第二が「限界費用の差異を利用したシグナリングメカニズム」である。

6. ハビトゥスによる選抜においてもパスワード改訂現象は生じているのではないか(仮説)

第一の流れとして、ハビトゥスによる選抜においては、パスワード改訂現象が生じているのではないかとわれわれとしては推測（仮説）してみたい。こう考えることで、上述のハビトゥスによる選抜の謎（の少なくとも一部）が解かれる、とおもわれる。説明しよう。

すなわち、もしハビトゥスによる選抜が、あるハビトゥスをもっている者たちに有利に・他のハビトゥスをもっている者たちに不利に、なるのだとしたら、たとえ無意識であっても、長期的には、下位ハビトゥス保持者たちも上位ハビトゥスを習得してしまう蓋然性がある（高い）だろう。それにたいするいわば対抗方略として、上位ハビトゥス自体が「変異」して、いわば下位ハビトゥス保持者による「おっかけ」を「ふりきる」メカニズムが進化したのではないかと考えてみたいのである。まさに、パスワード改訂メカニズムがここにも、生じているのではないかと。このようにかんがえられるのではないだろうか。

ブルデューが直接あつかった論件について、この仮説を支持する知見が存在するかどうかは知らない。しかし、われわれ日本人には、この仮説を支持すると思われる知見・現象が身近に存在する。すなわち、柳田國男が見いだした方言周圏現象である（柳田 1930）。

7. 方言周圏現象

周知のように、方言周圏現象とは、かたつむりなどの方言が、細長い日本列島において縞模様状に、すなわち（もし列島が細長くなかったと仮想すると）同心円状に分布しているという現象である。

そして、この方言の分布は、文化的中心地（すなわち、都）から、辺境にむけてある語の普及が伝播していったとみなされる。

とすると、（言語という）階層的ハビトゥスをめぐって、ここにおいてパスワード改訂現象が生じていたと推測できる蓋然性が高いのではないだろうか。

すなわち、文化的下位者たちは、文化的上位者たちのハビトゥス（ここでは言語）を、「おっかけ」的に習得したのだろう。しかし、その一方で、文化的上位者たちは、下位者たちの追撃（おっかけ）をふりきるように、新ハビトゥス（ここでは言語）を改訂していったのではないかと。言語は、比較的観察しやすいハビトゥスであるとおもう。がゆえに、ここではとくに言語をめぐって、パスワード改訂現象・その結果としての周圏現象、が観察されている。

しかし、それはじつはとくに言語にとどまらず多くの種類のハビトゥスにおいても平行して生じていたと私は見通している（もし、この「見通し」に対応する新事実が観察されれば、この視点の信憑性はなお増すだろう。どなたか、もしおこころあたりのある現象をご存じならご教示いただきたい）。

8. 「ベンツ」的シグナリング・メカニズム

ここで、一見すると上記とはすこしことなっているかのように見える視点から、ハビトゥスによる選抜論への疑問を再考してみよう。

うえでは、たとえ無意識であっても、例の暗黙の選抜過程において、上位階級のハビトゥス（育ちの良さを示すような習慣）が優遇されるのなら、なぜ、下位階級の者もそのハビトゥスを習得しないのか、と疑問点を提示した。

この疑問には、じつは、素朴な回答がありうるだろう。すなわち、下位階級の者（貧乏人）には、上位階級のハビトゥス（「上品」な習慣）を、習得するだけの「余裕がない」から、というものである。たとえば、「クラシック音楽、や、美術への、嗜好」によって、「お里が知れて」しまう事例を想起してみよう。

この回答は、じつはかなり正鵠を射ているとおもわれる。そしてまた、上記のパスワード改訂仮説と通底し、そしてまたそれを包含するものであるとおもわれる。説明しよう。

まず、すぐうえでのべた、「貧乏人には、金持ちハビトゥスを習得する余裕がない」という点が、情報経済学における「シグナリング」理論の典型的な実例（ケース）となっていることを確認しよう（Spence 1973）。

「余裕がない」と直観的にのべたが、厳密には、コストをかけて金持ちハビトゥスを習得しても、それに見合うだけのベネフィットをえられず、ペイしない、ということである。他方それにたいして、金持ちの方は、おやが同様なハビトゥスをもっている、とか、もともと金持ちなので習得（に、たとえ貧乏人と同じ金額がかかったとしても）のさいの貨幣の限界負効用が少ない、とかして、コストが相対的に貧乏人ほどかからない。あるいは、貧乏人をふりきることのベネフィットが大きい。こうして、都合、ベネフィットとコストの差し引きがペイする、ということである。

9. シグナリングメカニズムとしてのパスワード改訂

いうまでもないが、上記のシグナリングメカニズムをもたらす「ネタ」はなんでもいい。上記の関係をみたまものなら、なんでも OK である。クラシック音楽や美術への嗜好の多寡、あるいは、お茶・お花への通暁の多寡、さらには、（入試にはつかえないが）持っている車（ベンツ?!）や服装の値段など、つかえるものはさまざまだろう。

しかし、いうまでもなく、このシグナリングメカニズムは、（上位者階層にとっても）コストがかかってしまう。これは、上位者もコストをかけることで、下位者との一種の「がまんくらべ」をして、下位者をふりきるメカニズムであるからだ。

というわけで、この（とくに上位者階層にとっての）コストがすくなくればすくないほど、「性能の良い」シグナリングツールであるといえるだろう。

と、ここまでいえば、上述の「パスワード改訂」が、このような意味で比較的「性能の良い」（上位者にとってコストの低い）シグナリングツールであるといえるということはもはやいうまでもないだろう。

なにしろ、言語（とくに母語）には、習得上の臨界期が存在する。臨界期まえに、習得してしまえば、臨界期後にそれを「まねよう」としても、多大なコストがかかる。ただし、まったく変化させないでおけば、長期的には（何代もかかれば）「まね」されてしまうかもしれない。よって、「うちわ」のものにだけわかるような仕方で「改訂」しておく。そもそも、言語の体系は恣意的なものなので、この改訂は「行き届き」プロセスさえしっかり確保しておけば、比較的低いコストで可能となる、だろう。

10. パラサイト・シグナリングモデル、へむけて

さて、以上のように、現代ダーウィニズムから導出されたパスワード改訂仮説を、情報の経済学におけるシグナリング理論と接合させてみた。こうすることによってさらに、とくにすぐ上で触れたブルデューが論じた論件を再び想起することで、少しく興味深い視点を新たに提起できるように思う。いわば、「パラサイト・シグナリング」モデル、である。説明しよう。

ブルデューのハビトゥスによる選抜論において典型的にあつかわれるケースは、入試である。しかし、シグナリング理論をご存じの読者にはいうまでもないことだが、入試自身がそもそも、シグナリングの働きを持つものとして、機能しているのである。（以下数段落は、シグナリングモデルが既知の読者はとばしてください）。

ここに、能力が、「高」と「低」の二種類の学生がいたとしよう。学生は自分の能力の程度を知っているが、彼らを採用しようとする会社の方にはその情報が見えない、としよう（情報の非対称性）。

ここで会社側は、彼らを採用するさいに、「学歴」をシグナリングとして利用することが考えられる。ここでの眼目は二点ある。第一は、この学歴（教育）は、それ自体はなんら生産性の向上をもたらさなくていい、ということである。第二は、以下の帰結が生じるためには、その当該の教育（学歴）を獲得する限界費用が、能力「低」の学生の場合よりも、能力「高」の学生の場合の方が、低くなっているということが必要条件となっているということである。

いちばん、わかりやすい例としては、受験勉強にコスト（てまひま）がかかる「入試」をシグナリングにすればいい。ただし、その入試は、能力「高」の学生には比較して（限界的に）低いコストで、能力「低」の学生には比較的（限界的に）高いコストで、はじめて点数がのびるようなものであることが必要である。

このような前提条件（必要条件）のもとであるので、「いい会社」に入れる/入れない、にしても、「能力の高い学生には、合格のための受験勉強のコストを控除しても、なおその会社には入れることがペイし」、かつ、「能力の低い学生には、合格のための受験勉強のコストを

控除すると、その会社にはいってもペイしない」ような、「ちょうどよい合格水準」を設定することが可能になる。

このように、入試というものは、かなり本来的な機能のレベルで、もともとシグナリング機能を担っていた（場合が多い）といえそうである。

ところが、もともと（?）、このように受験生の本来の能力を測り取るために設計されているはずの入試が、ブルデューの事例においては受験生の別の属性（上では、階層に相関するハビトゥス）を「ふるいわけるといってはたらき「をも」してしまったのである。

このように、「本来の」属性峻別の機能として設計されたシグナリングツールが、べつの属性の峻別をもしてしまうということは、意外に多いように直観される。

このような「もともと開示させたい本来の、非対称情報」だけでなく、その峻別のいわば「副作用」として、「べつの非対称情報」をも峻別してしまう現象を、パラサイト・シグナリング現象（寄生シグナリング現象）と、呼んでみよう。

このようなパラサイトシグナリング現象は、この他にも、社会のさまざまな領域で生起しているように直観される。われわれのアプローチの今後の課題の一つとして、このようなパラサイト・シグナリング現象を、さまざまないままで気づかれていなかった論件において、発見することを目指したいとおもう。

また、それを関連したもう一つの課題がありうる。すなわち、いわゆるメカニズム・デザインの問題である。

通常情報の経済学において、メカニズムデザインというと、秘匿されている非対称情報を、誘因にそくした構造を設計することで、開示させ、より望ましい社会メカニズムをデザインする課題として考えられているだろう。

われわれの議論からは、それと非常に類似しているが厳密には若干ことなる課題を意識化することが可能になる、だろう。

すなわち、多くの場合、シグナリングメカニズムは、設計者が望む（非対称）情報のみならず、当事者の別の属性に即したふるい分け（フィルタリング）として機能してしまう（とわれわれは予想する）。いかなるメカニズムをデザインすることによって、望む属性によるフィルタリングをしつつ望まない属性によるフィルタリングを働かせないか、という課題である。

11. 二重(多重)シグナリングに対する、暗黙のご了解

以上のような視点を、さらにうがってかんがえると、さらに以下のようなシナリオをも構想してみることが可能になるとおもう。いわば、「暗黙の、二重のご了解」現象である。あるいは、「タテマエ、と、ホンネ、との出来レース」モデルである。説明しよう。

すなわち、まず、第一に、このような「パラサイトシグナリング」現象がしょうじているがゆえに、上位階層は、このメカニズムを支持（すくなくとも受容）していた、のではないかということである（「第一の」（一重の）暗黙のご了解）。

すなわち、なぜ、社会的な力をもっている上位階層がこのような「公平な」（そうであるが

ゆえに、絶対的に有利な自分のポジションにたいして相対的に不利な)メカニズムを受容しているのか。さまざまな理由がありうるだろうが、その大きな一つは、上記のパラサイト・シグナリングメカニズムによって、結果的に自分たちに有利な帰結がもたらされるからだ、とかんがえることができるだろう。(第一の、暗黙のご了解)。

ただし、この点は、すでに、ブルデューが、外見上の中立性・公平性がなぜ呼び込まれるのかについて論述していることで、實際上すでにのべられていたと解釈しうるかもしれない。

筆者としては、さらに「うがった」仮説をたててみたい。すなわち、すべての場合ではないにせよかなり多くの場合において、上記のような入試シグナリングのもつ、二重の機能(表のシグナリング、と、裏のシグナリング)を、じつは下位階層も無意識には了解しているのではないか、ということである。そして、相対的劣位者との優位者との「暗黙の交渉」によって、「名目ともに、とられるよりはマシ」として、いわば「名を取って、実をゆずっている」のではないか、ということである。

すなわち、「表の、「公平な」シグナリング」と「裏の、「優位者に有利な」シグナリング」の併存を、優位者と劣位者との、暗黙の合意、ないし、暗黙の妥協、いわば「手打ちの結果」として解釈してみたいのである。

こうして、「二重の、暗黙の、ご了解」として、二重のシグナリングメカニズムの社会的受容を解釈しうるのではないか、とかんがてみたいのである。

この第二の点は、ブルデューの視点とはかなり異なった視点(というか、まさに正反対の視点)であるように感じられる。がゆえに、これがすべての場合でないにせよ成立しているとする、本アプローチは、ブルデューの論圏の圏外へと脱したといえるだろう。が、いうまでもなく、この後者の仮説の実証は容易ではない。しかし、すくなくともある一つの場合には、ブルデュー的仮説と我々の仮説は、そこから予測される帰結を異にし、そうであるがゆえに、「実証的決着」がつけうる、と思われる。すなわち,,,,,

12. 社会変動期における、「あらかじめ約束されていた、暴露」

すなわち、ある種の社会変動期の場合において、である。筆者は、暗黙の二重のご了解を仮説した。すなわち、劣位者は、所与の状況のもとで、ないよりましな一種の「妥協」として、「名目的には公平な(多くの場合は、たんに名目的に公平であるだけでなく、実質上の「成り上がり」も「ある程度において」可能であろう)」しかし「実質的・暗黙的には、優位者有利」な、「二重シグナリングの組み合わせ」を受容(許容)していたのであった。

しかし、状況が変化すれば、劣位者によるこのような暗黙のシグナリングメカニズムの受容が、「公然化」する蓋然性・誘因が生じるだろう。いうまでもなく、劣位者たちが、優位者たちとのこのような社会関係から離脱する誘因をもちはじめたとき、である。

すなわち、「今(これまで)よりも、もっとヨリましな、選択肢」を劣位者が持つことによって、はじめて、劣位者たちは今までの関係を「虐げられていた」「不公平な」ものとして指弾する誘因をもつことになる、だろう。もし、このようなヨリましな選択肢がなかった場合、現状を、不公正なものとして指弾したとしても、みずからの自覚のもとでその不公正な関係

に甘んずる以外にかれらに道はない。

ここから、ブルデュー的仮定をとる場合、と、われわれのような仮定をとる場合での、「実証可能な差異」を導出することができるだろう。

すなわち、もしわれわれの仮説がただしければ、「劣位者の境遇が好転」したばあいに、いままでの選抜過程は、不公正であったという指弾が増大するだろう。

それにたいして、もしブルデューの仮説がただしければ、いままではソクな選抜過程であったのに、外見上の公正性にめくらましされていたのだが、より劣位者の境遇が悪化した場合に、それまでは閾値内であったがゆえに気づかれなかった自らの不利を感知し、不公正な選抜過程への指弾が生起するだろう。

このように、このようなケースに照準することで、われわれの仮説とブルデュー的仮説との「実証的決着」が可能になる、だろう。

13. 「ジェンダーバイアス」の指弾は、いつ高まる、か？

さらに、上記のような「素朴ブルデュー主義的仮説との訣別を可能にする分別可能な実証」の一スペシャルケース（特殊事例）として、いわゆる「ジェンダーバイアス」に関するちょっと非・常識的な検証仮説をたててみるのも、興味深いとおもう。

昨今のジェンダー論的文化分析・社会分析・テキスト分析の興隆についてはいうまでないだろう。これらには、さまざまなものがあるだろうが、強引に単純化してしまえば、もっともありがたいパタンのひとつは次のようなものだろう。

すなわち、いままでジェンダー（社会的性差）とは無関係ないし中立とみなされていた（というか正確には、なにもみなされていなかった）ような文化的社会的事象や諸テキストが、じつは、ジェンダー的にバイアスのかかったものであった、ということを暴露する、というものだろう。

これにたいしても、われわれの仮説的視点としては、これらの事象ならびにそれらに対するジェンダーバイアスとしての指弾それ自体（もちろんすべてではないが）を、例の「暗黙の二重のご了解」ならびに「そのお役ご免による崩壊（約束された暴露）」として解釈するわけである。

すなわち、当のジェンダーバイアスのかかった（つまり性別によって有利不利のある）文化的社会的事象を、上で述べてきた「二重のシグナリング」ツールとして解釈するわけである。「タテマエ上」はジェンダーに中立な（たとえば）能力主義的なふるい分け機能と、「暗黙の」ジェンダーふるい分け機能、という「オモテとウラの二重機能」をもったものと解釈するわけである。

そして、ジェンダー的な不利者たちも、多くの場合、このような「オモテとウラの二重機能」を暗黙で了解していた（オモテの中立的ふるい分けがあるだけ、「それさえもない、よりは、マシ」だから）。と、解釈するわけである。

注意しておくが、筆者は、なにも、ここでの言及にあたる事象が「すべて」、このように不利者によって暗黙に「ご了解」されていた、とは主張しない。しかし、すくなくならずは、こ

のように暗黙に了解されていたのではないかと仮説するのである。

もし、このような視点がなにほどであったにせよ、正鵠を射ていたとすると、そうでないとかんがえる対抗仮説に対して、実証的な決着可能な予測を導出することができるだろう。すなわち、

[予測]他の変数が同等であるとする、ジェンダー的な不利者の境遇が「好転」すればするほど、それまでジェンダー的に中立であったとされる文化的・社会的事象にたいして、「ジェンダーバイアスである」とする指弾、がなされる蓋然性がたかまるだろう、と。

14. おわりに

本稿は、現代ダーウィニズム、とくに、ダンバー的「方言変異」論が、人文社会諸科学にどのような洞察と通底性を与えてくれるかを、示そうとした。その結果、いろいろなトピックに浅く広くふれることになってしまった。この点、試論的な序説という本稿の位置づけからして、ある程度仕方がなかったかとおもう。とはいえ、読者にはいささか読みにくいペーパーとなってしまったかもしれない。今後は、ここで提起した個々の論点をさらに精緻に理論化し、実証へと接続させたい。

【文献】

- Bourdieu, Pierre, 1979. *La distinction : critique sociale du jugement* = 石井洋二郎訳. 1989. 『ディスタクシオン : 社会的判断力批判』. 新評論
- Bourdieu, Pierre, Passeron, Jean Claude, 1970. *La reproduction : elements pour une theorie du systeme d'enseignement* = 宮島喬訳. 1991. 『再生産 : 教育・社会・文化』 藤原書店
- Dunbar, R. I. M., 1996. *Grooming, gossip and the evolution of language* = 松浦俊輔, 服部清美訳. 1998. 『ことばの起源 : 猿の毛づくろい, 人のゴシップ』. 青土社
- Harris, Judith Rich, 1998. *The nurture assumption : why children turn out the way they do* = 石田理恵訳. 2000. 『子育ての大誤解 : 子どもの性格を決定するものは何か』. 早川書房
- Spence, A Michael, 1973. "Job Market Signaling," *The Quarterly Journal of Economics*, Vol. 87 (3) pp. 355-74
- 柳田國男. 1930. 『蝸牛考』. 刀江書院

Shiro Horiuchi

A competition model within and across groups explaining the contrast between the societies of chimpanzees and bonobos

Received: 7 November 2003 / Accepted: 25 February 2004 / Published online: 21 April 2004
© The Society of Population Ecology and Springer-Verlag Tokyo 2004

Abstract Two apes species, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), show a peculiar contrast in their social structure. The contrast cannot be explained simply by the differences in their ecological conditions, as has previously been suggested. This paper presents a mathematical model to interpret the contrast, through a competition game within and across groups. Two alternative strategies, dispersal and association, are applied to the model, in order to represent within-group-competition (WGC). Two more alternative strategies, Hawk and Dove, are also applied to the model, in order to represent between-group-competition (BGC). Mathematical analysis is executed for the frequencies of three mixed-strategists: dispersal-Hawk, dispersal-Dove and association. The population dominated by dispersal-Hawk and dispersal-Dove strategists, or by association strategists, is assumed to characterize the social structure of chimpanzees or bonobos, respectively. The model predicts that each social structure can be evolutionarily stable under the same ecological conditions, and will not easily recover once it has been replaced by the other. This stability is through a bi-directional influence: the influence of BGC on WGC and vice versa. The model suggests that the possibility of fatal violence within a group strengthens the robustness of each social structure.

Keywords Hawk–Dove game · Dispersal strategy · Association strategy · Within group competitions · Between group competitions · Robust social structure

S. Horiuchi
Laboratory of Human Evolution Studies,
Department of Zoology, Graduate School of Science,
Kyoto University, Kyoto, Japan

Present address: S. Horiuchi
Division of Behavioral Science, Faculty of Arts and Letters,
Tohoku University, Kawauchi,
Aoba-ku, Sendai 980-0862, Japan
E-mail: horiuchi@sal.tohoku.ac.jp
Tel.: +81-22-217-6037
Fax: +81-22-217-5972

Introduction

The two species of the genus *Pan*, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), have been studied for several decades by many researchers (Goodall 1986; Nishida 1990; Kano 1992). Both form male-philopatric groups in fission–fusion societies (Nishida 1968; Kano 1982). Their social structures, however, show a peculiar contrast (Nishida and Hiraiwa-Hasegawa 1987; Wrangham and Peterson 1996). Roughly speaking, chimpanzees are characterized as violent apes. Their competitions are aggressive both within (Fawcett and Muhumuza 2000) and across groups (Nishida et al. 1985). By contrast, bonobos are characterized as peaceful apes. Their competitions are relaxed both within (Kuroda 1980) and across groups (Idani 1990). In particular, infanticides and lethal raids have been reported several times among chimpanzees but never among bonobos (Goodall 1986; Nishida 1990; Kano 1992).

The contrast is often attributed to the stability of female association which, according to one hypothesis, is caused by the food distribution pattern (Wrangham 1986). It is said that the food patch of chimpanzees is small and seasonal, and females cannot associate with one another because of their strong competition for food resources within a group (Wrangham 1979). Therefore, multiple males occasionally patrol the boundary of their group to guard dispersing females against males from neighboring groups. Hence, when males or females feed alone they occasionally suffer lethal raids from patrolling males from neighboring groups (Wrangham 1999). On the other hand, it is said that the food patch of bonobos is large and not seasonal, and females can always associate with one another (White 1998). Therefore, females as well as males usually join a large, mixed-sex party, and do not feed alone. Hence, they only rarely suffer lethal raids by patrolling males from neighboring groups (Wrangham 1999). In other words, it is assumed that the food distribution pattern influ-

ences within-group-competition (WGC, van Schaik 1983), which in turn influences between-group-competition (BGC, Wrangham 1980), in the two ape species.

The hypothesis of Wrangham (1986), however, cannot fully explain the contrast between the social structures of the two ape species. Different chimpanzee populations inhabit various environments, from rich to poor. They nevertheless maintain a more or less similar pattern of social structure, within and across groups (Gombe, Goodall 1986; Mahale, Nishida 1990; Kibale, Wrangham 1999; Budongo, Newton-Fisher 1999a; Fawcett and Muhumuza 2000; Tai, Boesch and Boesch 2000). Furthermore, the environments inhabited by bonobos are not necessarily richer than those inhabited by chimpanzees (Chapman et al. 1994). Each social structure seems to be robust against changes in ecological conditions.

One could hypothesize that BGC and WGC influence each other. Hence, each social structure has robustness against changes in ecological conditions. I present here a simple game model (Maynard-Smith 1982) that considers mutual interaction of WGC and BGC, where each social structure has robustness against changes in ecological conditions. A game model is useful as a heuristic device to elucidate the condition causing the robustness of social structure (Matsumura and Kobayashi 1998).

Assumption of the model

Both chimpanzee and bonobo individuals are assumed to be variants of the same game model. All individuals are assumed to use one of two applied strategies, dispersal and association, which represent their fission–fusion societies. Dispersal strategists feed alone. Association strategists follow their group mates. Because of WGC, association strategists are assumed to suffer cost D_S ($D_S > 0$). Furthermore, they suffer additive cost D_C ($D_C > 0$) proportional to the frequency of Hawk strategists (see later sections) within their group: Hawk strategists are assumed to inflict violence on association strategists.

When two individuals from different groups see each other, they are assumed to have a battle. The battle follows a modified Hawk–Dove game (Maynard-Smith 1982) which represents BGC. Their group-mates only support the two attendants if the attendants are association strategists. Association strategists suffer WGC costs, but gain BGC benefits, since they thus avoid being involved in battles alone across groups.

There are three battle patterns possible across groups. If both the combatants are dispersal strategists, the battle follows the usual Hawk–Dove game (Fig. 1a); the winner gains benefit V , and the loser suffers no cost if it is a Dove strategist or suffers cost C if it is a Hawk strategist. For simplicity, I assume C is larger than V ($C > V > 0$). If one combatant is an association strategist and the other is a dispersal strategist, the former always defeats the latter, via the support of its group-mates

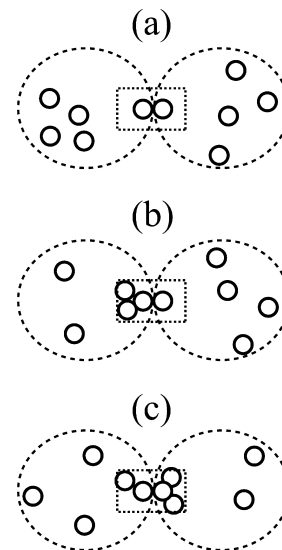


Fig. 1 Scheme of a battle across groups in three patterns. The *two dotted circles* represent two neighboring groups: each of them contains its members as *solid circles*. The *dotted square* contains individuals who are attending a battle. **a** Both the attendants are dispersal strategists. They execute the battle with no supporters, and the battle follows the Hawk–Dove game. **b** One attendant is a dispersal strategist and the other is an association strategist. The latter always defeats the former with the support of its group-mates. **c** Both the attendants are association strategists. The two attendants, with the support of their group-mates, execute the battle in a Dove-like manner

(Fig. 1b). The former gains benefit V , and the latter suffers no cost if it is a Dove strategist or suffers cost C if it is a Hawk strategist. If both the combatants are association strategists, they execute the battle in a Dove-like manner (Fig. 1c). Both gain benefit $V/2$ and suffer no cost. The group-mates who support the combatants are assumed not to gain benefit or suffer cost, for the sake of simplicity of the model.

The mathematical model has four parameters: D_S , D_C , V and C . This paper considers the value of D_S , rather than that of V , as representing the ecological condition for a population. High and low values of D_S mean poor and rich ecological conditions, respectively. Ecological conditions are assumed to change continuously from rich to poor or vice versa. The values of D_C and C represent severity of WGC and BGC, respectively.

Mathematical analysis is executed for the frequencies of three mixed-strategists: dispersal-Hawk (DH), dispersal-Dove (DD) and association (AS). Dispersal strategists can be Hawk or Dove strategists at battles between groups, since they behave independently from their group-mates. On the other hand, association strategists follow collective behavior at battles between groups, as described above. Accordingly, the three mixed strategies are possible for individuals. The frequencies of the three mixed-strategists are denoted as x , y and z ($0 \leq x, y, z \leq 1$, $x + y + z = 1$), respectively. Now $(x, y, z)^T$ and $(X, Y, Z)^T$ are the vectors of relative

frequencies of each mixed-strategist in the focal group and the neighboring group. With the dimensionless parameters defined above, I can describe the profit function of each mixed-strategist as follows.

$$\begin{pmatrix} W_{DH} \\ W_{DD} \\ W_{AS} \end{pmatrix} = - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ D_C + D_S & D_S & D_S \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{bmatrix} (V-C)/2 & V & -C \\ 0 & V/2 & 0 \\ V & V & V/2 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

All groups are assumed to have similar sizes. Therefore, it is plausible to assume that they have the same frequencies of the three mixed-strategists. Now substitute $(X, Y, Z)^T = (x, y, z)^T$ and $z = 1-x-y$. The profit functions of the three mixed-strategists are equivalent to

$$\begin{aligned} W_{DH} &= x(V+C)/2 + y(V+C) - C \\ W_{DD} &= yV/2 \\ W_{AS} &= x(V/2 - D_C) + yV/2 + V/2 - D_S. \end{aligned} \quad (1)$$

Fig. 2a-f Dynamics of x and y , representing the frequency of DH and DD, respectively (the frequency of AS is $1-x-y$). Two isoclines, Eqs. 3a and b, are represented as *dotted lines* and *broken lines*, respectively, on (x, y) space. (x, y) converge to each of the three points P1: $(0, 0)$; P2: $[(2D_S - V)/(V - D_C), \{V(V + 3C) - 2(V + C)D_S - 4CD_C\}/\{(V + 2C)(V - 2D_C)\}]$; or P3: $(V/C, 1 - V/C)$, depending on the values of D_C and D_S . **a** When $D_C < V/2$ and $D_S < V/2$, (x, y) converge to P1. **b** When $D_C < V/2$ and $V/2 < D_S < V(C + V - 2D_C)/(2C)$, (x, y) converge to P2. **c** When $D_C < V/2$ and $D_S > V(C + V - 2D_C)/(2C)$, (x, y) converge to P3. **d** When $D_C > V/2$ and $D_S < V(C + V - 2D_C)/(2C)$, (x, y) converge to P1. **e** When $D_C > V/2$ and $V(C + V - 2D_C)/(2C) < D_S < V/2$, (x, y) converge to P1 or P3, depending on their incipient frequencies. **f** When $D_C > V/2$ and $D_S > V/2$, (x, y) converge to P3

The population dominated by DH and DD, or by AS, is assumed to represent the social structure of chimpanzees or bonobos, respectively.

Results of the model

I can assume an evolutionary dynamics of x and y as,

$$x' = xW_{HD}/\{xW_{HD} + yW_{DD} + (1-x-y)W_{AS}\}, \quad (2a)$$

$$y' = yW_{DD}/\{xW_{HD} + yW_{DD} + (1-x-y)W_{AS}\}. \quad (2b)$$

The two equations, $x' > x$ and $y' > y$, are respectively equivalent to

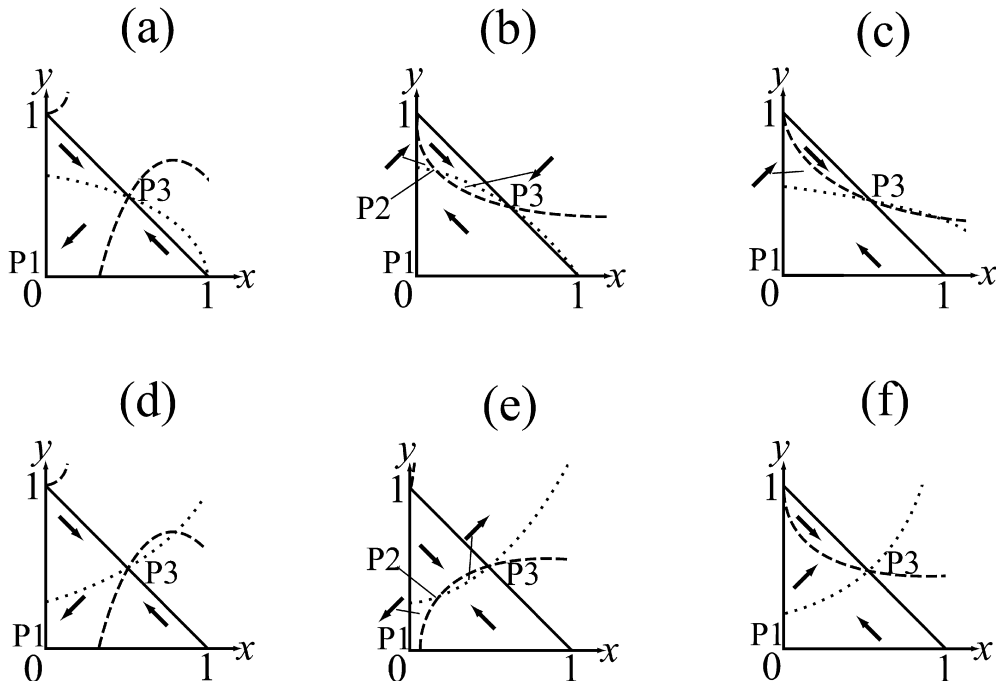
$$\begin{aligned} 2(D_S - C) - V + x\{V + 3C + 2(D_C - D_S)\} \\ + 2y(V + C - D_S) - x^2(C + 2D_C) - 2xy(C + D_C) \\ > 0, \end{aligned} \quad (3a)$$

$$\begin{aligned} 2D_S - V + 2x(C - D_S + D_C) + y(V - 2D_S) \\ - x^2(C + 2D_C) - 2xy(C + D_C) > 0. \end{aligned} \quad (3b)$$

Fig. 2 represents the two isoclines (see Appendix 1). By following the dynamics of x and y , I can investigate whether their equilibrium points are QUERY ESSs (Maynard-Smith 1982) or not. The dynamics depends on the values of D_C and D_S . If $D_C < V/2$, three ESSs appear, depending on the value of D_S . When $D_S < V/2$ (Fig. 2a), the Evolutionary Stable State (ESS) is

$$P1 = (0, 0).$$

When $V/2 < D_S < V(C + V - 2D_C)/(2C)$ (Fig. 2b), the ESS is



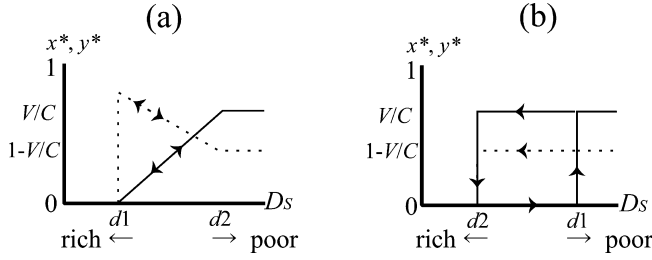


Fig. 3a, b Changes in x^* (solid line) and y^* (dotted line), representing the frequency of DH and DD at ESS, respectively, along the value of D_S (the frequency of AS at ESS is $1-x^*-y^*$). As the value of D_S is high, ecological condition is assumed to be poor, and the sum of x^* and y^* becomes small. d_1 and d_2 in the figure are $V/2$ and $V(C+V-2D_C)/(2C)$, respectively. x^* and y^* change differently, depending on the value of D_C . **a** $D_C < V/2$. (x^* , y^*) are P1, P2, or P3, when $D_S < d_1$, $d_1 < D_S < d_2$, or $d_2 < D_S$, respectively. x^* or y^* can change bi-directionally at any point of D_S within the region of $d_1 < D_S < d_2$. **b** $D_C > V/2$. (x^* , y^*) are P1, P1 or P3, or P3, when $D_S < d_2$, $d_2 < D_S < d_1$, or $d_1 < D_S$, respectively. Once the value of D_S has increased over the value of d_1 , so that (x^* , y^*) becomes P3, x^* or y^* will not decrease until the value of D_S decreases below the value of d_2 . Once the value of D_S has decreased below the value of d_2 , so that (x^* , y^*) becomes P1, x^* or y^* will not increase until the value of D_S increases over the value of d_1

$$P2 = \left(\frac{2D_S - V}{V - 2D_C}, \frac{V(V + 3C) - 2(V + C)D_S - 4CD_C}{(V + 2C)(V - 2D_C)} \right).$$

Local stability is proved with an analysis of a Jacobian matrix (see Appendix 2). When $D_S > V(C+V-2D_C)/(2C)$ (Fig. 2c), the ESS is

$$P3 = (V/C, 1 - V/C).$$

On the other hand, if $D_C > V/2$, two ESSs appear, depending on the value of D_S . When $D_S < V(C+V-2D_C)/(2C)$ (Fig. 2d), the ESS is

$$P1 = (0, 0).$$

When $V(C+V-2D_C)/(2C) < D_S < V/2$ (Fig. 2e), two bi-stable ESSs appear. They are

$$P1 = (0, 0) \text{ and } P3 = (V/C, 1 - V/C).$$

When $D_S > V/2$ (Fig. 2f), the ESS is

$$P3 = (V/C, 1 - V/C).$$

Discussion

Figure 3 represents the frequencies of DH and DD at ESS, along the value of D_S . Each frequency is represented as x^* and y^* , respectively. If $D_C < V/2$, (x^* , y^*) can be each of the three states of P1, P2 or P3 (Fig. 3a), corresponding to the case of Fig. 2a, b, or c, respectively. The frequencies of any mixed-strategists change gradually along the value of D_S . That means the social structure of chimpanzees or bonobos replaces itself

continuously, through changes in ecological conditions, as discussed by Wrangham (1986). If $D_C > V/2$, (x^* , y^*) can be each of the two states of P1 or P3 (Fig. 3b), corresponding to the case of Fig. 2d and e, or 2e and f, respectively. The frequencies of any mixed-strategist neither increase nor decrease without drastic changes in the value of D_S . That means that once the social structure of bonobos replaces that of chimpanzees, or vice versa, the original social structure will not recover until drastic reverse changes in ecological conditions occur. Hence, the present model shows an existence of the robustness of each social structure against changes in ecological conditions.

As the value of D_C becomes higher, the bi-stable region along the value of D_S , or $V(C+V-2D_C)/(2C) < D_S < V/2$, becomes wider. Thus, the possibility of fatal violence within a group, such as infanticide, is predicted to cause the robustness of their social structures against changes in ecological conditions (Kano 2001). In fact, party sizes of chimpanzees do not increase flexibly even if the sizes of the food patch or the numbers of estrous females increase (Newton-Fisher 1999b). Chimpanzees could not change their social behaviors flexibly according to changes in ecological conditions, unlike their feeding behaviors (Yamakoshi 1998).

Each of the two social structures is maintained as follows. Once DH and DD fill the population at the frequency of V/C and $1-V/C$, respectively, only two individuals with no supporters execute most of the battles across groups. Since DH and DD do not suffer collective attacks from AS, their frequencies are maintained at a high. Since DH gave high costs to AS, AS cannot invade the population. On the other hand, once AS fill the population, other group-mates always support combatants between groups. Since DH and DD always lose battles between groups, they cannot invade the population. Since AS do not suffer violence from DH within a group, their frequencies are maintained at a high. In this way, the manner of BGC stabilizes that of WGC and vice versa.

The model presented in this paper may oversimplify the social structures of the two ape species. For example, it does not distinguish the two sexes explicitly. In fact, males are more cohesive than females in chimpanzees, and the contrary is the case in bonobos (White and Chapman 1994). In chimpanzees, a party composed of multiple males (not a mixed-sex group) commits lethal raids against males of the neighboring group during boundary patrol (Nishida and Hiraiwa-Hasegawa 1987). Our purpose, however, is not to construct the two social structures completely but to find their robustness. Through the robustness identified by the model, the two social structures might have evolved through differentiation of the two sexes (Hemelrijk 2002). The model also assumed that two neighboring groups are similar sizes. In fact, a group of larger size would destroy the smaller-sized group; the males would be killed and the females recruited from the losing group to the winning group (Nishida et al. 1985). Since extinctions of groups are rare

in chimpanzees, I believe my model is enough to consider the contrast between the social structures of the two ape species.

The mathematical model would appear plausible if it could explain a similar contrast in other species. Spider monkeys (*Ateles* sp.) and woolly monkeys (*Lagothrix lagotricha*) form male-philopatric and fission–fusion societies (Symington 1990). The two species are widespread and sympatric in the western Amazon region and are similarly frugivorous (Iwanaga and Ferrari 2001). However, they show a peculiar contrast in social structure: party size is small and relationships between groups are agonistic in spider monkeys, while party size is large and relationships between groups are peaceful in woolly monkeys (Nishimura 1994; Strier 1994). Some populations of woolly monkeys try to maintain the cohesiveness of their group by feeding on insects, even when the food quantity and quality is poor (Di Fiore and Rodman 2001). That is, two species of atelines, geographically and genetically distant from *Pan*, show us a similar contrast as that found in *Pan*. The same model presented here could also explain the contrast of those two social structures.

Chimpanzees and bonobos have been used as the models for reconstructing hominid social structure (Wrangham and Peterson 1996). We do not know exactly which type of social structure the early hominids had, whether similar to that of chimpanzees or that of bonobos. Chimpanzees are assumed to have acquired social structures suited for poor ecological conditions when they scattered in various environments. Although some populations now inhabit as rich environments as bonobos, they appear to maintain their robust social structure suited for poor environmental conditions. Bonobos may also maintain their social structure, even if the ecological conditions of their habitats become as poor as those inhabited by chimpanzees. Hence, our model suggests that either type is possible as the prototype of the early hominids' social structure. In any possible habitat, they could have kept their social structure stable against changes in ecological conditions.

Acknowledgments I thank T. Nishida, J. Yamagiwa, S. Suzuki, and other members of the Laboratory of Human Evolutionary Studies, Kyoto University for their valuable advice and discussion. I thank K. Adachi, H. Takasaki, and G. Yamakoshi for carefully reading early drafts and for their critical comments. I also thank A. Nishimura for permission to read his unpublished paper. I also thank one anonymous referee for improvement in my manuscript. This research was financially supported by the MEXT Grant-in-Aid for the 21st COE Program, Biodiversity (A14).

Appendix 1

The two isoclines cross each other only at P2 and P3. P2 is within the region of $0 < x$, $0 < y$ and $x + y < 1$, when $D_C < V/2$ and $V/2 < D_S < V(C + V - 2D_C)/(2C)$, or when $D_C > V/2$ and $V(C + V - 2D_C)/(2C) < D_S < V/2$. P3

is always on the line of $x + y = 1$. To analyze the dynamics of x and y , it is enough to investigate the six cases of (a) $D_C < V/2$ and $D_S < V/2$, (b) $D_C < V/2$ and $V/2 < D_S < V(C + V - 2D_C)/(2C)$, (c) $D_C < V/2$ and $V(C + V - 2D_C)/(2C) < D_S$, (d) $D_C > V/2$ and $D_S < V(C + V - 2D_C)/(2C)$, (e) $D_C > V/2$ and $V(C + V - 2D_C)/(2C) < D_S < V/2$, and (f) $D_C > V/2$ and $V/2 < D_S$.

Now transform Eqs. 3a and b into

$$2y\{(V + C - D_S) - (C + D_C)x\} > (1 - x)\{(2C + V - 2D_S) - (C + 2D_C)x\}, \quad (4)$$

$$y\{(V - 2D_S) - 2(C + D_C)x\} > (C + 2D_C)(x - X1)(x - X2), \quad (5)$$

respectively, where $X1 < X2$ and $X1, X2 = [C + D_C - D_S \pm \sqrt{\{(C - V)(C + 2D_C) + (D_C + D_S)^2\}}]/(C + 2D_C)$. The shape of Eq. 4 depends on the relationship between values of 1, $(V + C - D_S)/(C + D_C)$, and $(2C + V - 2D_S)/(C + 2D_C)$. The shape of Eq. 5 depends on the relationships between values of $(V - 2D_S)/[2(C + D_C)]$, $X1$, and $X2$.

In the case of (a), simple algebra shows that $1 < (V + C - D_S)/(C + D_C) < (2C + V - 2D_S)/(C + 2D_C)$ and $(V - 2D_S)/[2(C + D_C)] < X1 < X2$. Therefore, I can illustrate the two isoclines as Fig. 2a. The ESS is P1. In the same way, I can illustrate Fig. 2b–f to find the ESS in each case.

Appendix 2

To analyze the local stability of the system at P2, I derive the Jacobian matrix. Hereafter, it is assumed that $D_C < V/2$ and $V/2 < D_S < V(C + V - 2D_C)/(2C)$. Now rewrite Eqs. 2a and b as $x' = F(x, y)$ and $y' = G(x, y)$. The Jacobian matrix is

$$\begin{bmatrix} \partial F(x, y)/\partial x & \partial F(x, y)/\partial y \\ \partial G(x, y)/\partial x & \partial G(x, y)/\partial y \end{bmatrix}.$$

Now substitute $x = (2D_S - V)/(V - 2D_C)$ and $y = [V(V + 3C) - 2(V + C)D_S - 4CD_C]/(V + 2C)(V - 2D_C)$. The characteristic equation of this matrix is

$$\begin{aligned} r^2 V^2 \{V^2 + 3CV - 4CD_C - 2(C + V)D_S\} \\ - 2rV \{2CD_S^2 + (2VD_C - 2V^2 - 3CV - 2C^2)D_S \\ - (4C + V)VD_C + V(V + C)(V + 2C)\} \\ + 4CD_S \{(V + C)V - 2VD_C - 2CD_S\} = 0, \end{aligned}$$

where the eigenvalues are represented as r . Now symbolize the left side of this equation as $H(r)$. $H(r)$ has its minimum value at its axis, since the coefficient of r^2 is positive. The axis is

$$\begin{aligned} \{2CD_S^2 + (2VD_C - 2V^2 - 3CV - 2C^2)D_S - (4CV + V^2)D_C \\ + V^3 + 3V^2C + VC^2\}/V \{V^2 + 3CV - 4CD_C - 2(C + V)D_S\}. \end{aligned}$$

Simple algebra shows that the axis is larger than 0 and smaller than 1, and that $H(0) > 0$, $H(1) > 0$. Therefore, the absolute values of the eigenvalues are smaller than 1, as long as they are real numbers. The discriminant of the characteristic equation is $(2D_S - V)(D_C^2 A + D_C B + E)$, where

$$A = 2D_S V^2 - V(4C + V)^2,$$

$$B = 4D_S^2 VC - 2D_S(V + 2C)(2V^2 + 3VC + 4C^2) + 2V(V + 4C)(V^2 + 3(VC + C^2)),$$

$$E = 2D_S^3 C^2 - D_S^2 C(4V^2 + 13VC + 12C^2) + 2D_S(V^4 + 5V^3 C + 10V^2 C^2 + 9VC^3 + C^4) - V(V^2 + 3VC + C^2)^2.$$

Now consider the equation $I(D_C) = D_C^2 A + D_C B + E$. Simple algebra shows $A < 0$ and $0 < -B/(2A) < V/2$. Therefore, $I(D_C)$ has its minimum values at $D_C = 0$ or $D_C = V/2$. Simple algebra shows $I(0) > 0$ and $I(V/2) > 0$. Therefore, the discriminant is positive. Accordingly, it was shown that the absolute values of the eigenvalues are smaller than 1, and that P2 is locally stable.

References

- Boesch C, Boesch H (2000) The chimpanzees of the Taï forest: behavioral ecology and evolution. Oxford University Press, Oxford
- Chapman CA, White FJ, Wrangham RW (1994) Party size in chimpanzees and bonobos: a reevaluation of theory based on two similarly forested sites. In: Wrangham RW (ed) Chimpanzee culture. Harvard University Press, Cambridge, pp 41–57
- Di Fiore A, Rodman PS (2001) Time allocation patterns of lowland woolly monkeys (*Lagothrix lagotricha poeppigii*) in a neotropical Terra Firma forest. *Int J Primatol* 22:449–480
- Fawcett K, Muhumuza G (2000) Death of a wild chimpanzee community member: possible outcome of intense sexual competition. *Am J Primatol* 51:243–247
- Goodall J (1986) The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge
- Hemelrijk CK (2002) Despotic societies, sexual attraction and the emergence of male ‘tolerance’: an agent-based model. *Behaviour* 139:729–747
- Idani G (1990) Relations between unit-groups of bonobos at Wamba, Zaire: encounters and temporary fusions. *Afr Stud Mon* 11:153–186
- Iwanaga S, Ferrari SF (2001) Party size and diet of syntopic Atelids (*Ateles chamek* and *Lagothrix cana*) in Southwestern Brazilian Amazonia. *Folia Primatol* 72:217–227
- Kano T (1982) The social group of pygmy chimpanzees (*Pan paniscus*) of Wamba. *Primates* 23:171–188
- Kano T (1992) The last ape: pygmy chimpanzee behavior and ecology. Princeton University Press, Princeton
- Kano T (2001) Counter strategies against potential infanticide?: reconsideration of social characteristics of *Pan paniscus*. *Prim Res* 17:223–242 (in Japanese with English summary)
- Kuroda S (1980) Social behavior of the pygmy chimpanzees. *Primates* 21:181–197
- Matsumura S, Kobayashi T (1998) A game model for dominance relations among group-living animals. *Behav Ecol Sociobiol* 42:77–84
- Maynard-Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge
- Newton-Fisher NE (1999a) Infant killers of Budongo. *Folia Primatol* 70:168–169
- Newton-Fisher NE (1999b) Association by male chimpanzees: a social tactic? *Behaviour* 136:705–730
- Nishida T (1968) The social group of wild chimpanzees in the Mahali Mountains. *Primates* 9:167–224
- Nishida T (ed) (1990) The chimpanzees of the Mahale Mountains: sexual and life-history strategies. University of Tokyo Press, Tokyo
- Nishida T, Hiraiwa-Hasegawa M (1987) Chimpanzees and bonobos: cooperative relationships among males. In: Smuts BB, Cheney DL, Seyfarth RM, Wrangham WR, Struhsaker TT (eds) Primate societies. University of Chicago Press, Chicago, pp 165–177
- Nishida T, Hiraiwa-Hasegawa M, Hasegawa T, Takahata Y (1985) Group extinction and female transfer in wild chimpanzees in the Mahale National Park, Tanzania. *Z Tierpsychol* 67:284–301
- Nishimura A (1994) Social interaction patterns of woolly monkeys (*Lagothrix lagotricha*): a comparison among the Atelines. *Sci Eng Rev Dos Univ* 35:235–253
- van Schaik CP (1983) Why are diurnal primates living in groups? *Behaviour* 87:120–144
- Strier KB (1994) Brotherhood among Atelines: kinship, affiliation and competition. *Behaviour* 130:151–167
- Symington MM (1990) Fission–fusion social organization in *Ateles* and *Pan*. *Int J Primatol* 11:47–61
- White FJ (1998) Seasonality and socioecology: the importance of variation in fruit abundance to bonobo sociality. *Int J Primatol* 19:1013–1027
- White FJ, Chapman CA (1994) Contrasting chimpanzees and bonobos: nearest neighbor distances and choices. *Folia Primatol* 63:181–191
- Wrangham RW (1979) Sex differences in chimpanzee dispersion. In: Hamburg D, McCown ER (eds) The great apes. Benjamin/Cummings, Menlo Park, pp 13–53
- Wrangham RW (1980) An ecological model of female-bonded primate groups. *Behaviour* 75:262–300
- Wrangham RW (1986) Ecology and social relationships in two species of chimpanzee. In: Rubenstein DI, Wrangham RW (eds) Ecological aspects of social evolution: birds and mammals. Princeton University Press, Princeton, pp 352–378
- Wrangham RW (1999) Evolution of coalitionary killing. *Y Phys Anthropol* 42:1–30
- Wrangham RW, Peterson D (1996) Demonic males. Houghton Mifflin, New York
- Yamakoshi G (1998) Dietary responses to fruit scarcity of wild chimpanzees at Bossou, Guinea: possible implications for ecological importance of tool use. *Am J Phys Anthropol* 106:283–295

不確実性を伴う意思決定の非合理性に対する合理的説明

河野 敬雄 (広島修道大学)

要旨

決定論的世界観は、現実の力学的世界がニュートンに始まる微分方程式モデルを解くことによって決定論的に予測できる、という信念に裏付けられている。たとえ、現実にはあり得ないことが明白であっても綿 1 グラムと鉄 1 グラムを地上で落下させた場合、理想状態では同じ速度で落下する、と信じて疑わない。にもかかわらず不確実な予測を伴う場合、大数の法則を考慮したり、ベイズの定理を計算して予測に応用しよう、という態度は希薄である。逆に人間の直感によって得られる解が如何に不合理であるか、という研究は数多く行われている。本稿では、その中で、Boudon が主観的合理性を論じる際に引用してある例 ([10]) とゲーム理論の分野で知られているムカデ・ゲームについて、コルモゴロフ (1933, [22]) によって確立された純粋数学としての公理的確率論の公理を前提に、新しい合理的説明モデルを提案する。

キーワード：合理的意思決定、ムカデゲーム、主観確率

§0. 序

直感的な「確率」を実証と照らし合わせて検証しようという試みは結構古い歴史を持っている。今、3つのサイコロを同時に振って出た目の和が 9 である確率と 10 である確率を比較しよう。目の和が 9 となる場合は

$$(1, 2, 6), (1, 3, 5), (1, 4, 4), (2, 2, 5), (2, 3, 4), (3, 3, 3)$$

の 6 通りであるのに対して、目の和が 10 となる場合は

$$(1, 3, 6), (1, 4, 5), (2, 2, 6), (2, 3, 5), (2, 4, 4), (3, 3, 4)$$

であり、同じく 6 通りである。にもかかわらずイタリアの貴族 (= 賭博師?) による経験では目の和が 10 となる場合の方がわずかに起こりやすいことが知られていたようである。上記のように場合の数は同じであるから実現頻度はほぼ五分五分であろう、という素朴な推論は正しくない、何故か? と疑問に思い、当時最高の学者であったガリレオ・ガリレイ (1564-1642) に問いただした、ということ自体、まず評価されてよいことである。残念ながら筆者が知る限り、日本または中国の王侯貴族、知識階級の間で、このように数理的推論と事実との食い違いをさらに突き詰めて考察、探求しようとした例はないように思われる。

さて、ガリレオの解答はよく知られているように、たとえ見た目には区別がつかないサイコロであっても区別して場合の数を勘定せよ、というものであった。その結果は、3つの目の和が 9 である場合の数は 25、3つの目の和が 10 である場合の数は 27 となって、わずかではあるが 3つの目の和が 10 である場合の方が有利であり経験的事実をよく説明している、というもので

あった。経験的事実の集積，直感的推論，その検証，より深い考察に基づく数理モデルの提出，その検証というこの一連のプロセスは現代の（少なくとも自然科学の）研究においても基本的に何の変更を加えることなく通用する研究態度である。

しかしながら，不確実さを伴う判断ないし選択を強いられた場合「確率」の値がはっきりと指定されている場合でも，必ずしも「合理的」推論が出来ず，直感と合わない，といって実証もできそうにない例というものも古くから知られている。そのひとつがいわゆるペテルスブルグ問題（ペテルスブルグのパラドックス）である。

いま，A が公平な硬貨を表が出るまで投げ続けて， n 回目に初めて表が出たとするとその時点で A は 2^n のドルもらえる，という賭けに誘われたとする。このとき，A は何ドルまでなら支払ってこの賭けをするのが有利であろうか，という問題である。当然，A の受け取る賞金の期待値以下であれば掛け金を払って賭けをする方が有利である。ところが，期待値は実現確率 \times そのとき受け取る金額の合計であるから，

$$\text{このゲームで受け取る賞金の期待値} = \sum_{n=1}^{\infty} (2^{-n} \times 2^n) = +\infty$$

となつて，期待値は無量大，つまりいくら支払ってもこの賭けは A にとって有利なのである。しかし，多くの人は 10 ドルでもこの賭けのために支払おうとはしない，と言われている。何故か，とうことはこれまた古くから説明が試みられている。例えば，ダニエル・ベルヌーイ (1700-1752) は人間の評価（期待値）は金額が大きくなるに従って額面よりも小さくなる（効用逡減の法則を仮定した効用関数を導入することに相当する），として金額の対数をとって平均することを提案している。この場合，期待値が有限（きわめて小さい値）になることは確かであるが，客観的な実証データが得られるようには思われないので検証しようがない。

数学的には級数のいわゆる収束のオーダー（この場合は発散のオーダー）の問題のように思われる。つまり，2 百万ドル以上を得るためには 20 回以上裏が出続ける必要があるが，初めて表が出る平均回数はわずかに 2 回なのである。10 回裏が続けて出た場合はその時点で 2^{10} のドルを支払って賭けをやめて貰う，という変更をしても現実的にはさしたる変更ではないと感じるであろうが，この場合の期待値は 12 ドルである。数学的実証性から言えば，例えば 1 秒間に 100 億回の賭けをして，それを 100 億年続けた場合には，高い確率で十分大きな算術平均の値が得られるであろうが，現実的であるとは誰しも思わないであろう。

ペテルスブルグのパラドックスに関してはその後多くの人（ただし，数学者以外の人が多いようであるが）が論じているが筆者を納得させるような説明は見当たらない。

確率評価のパラドックスとして心理学の分野の実証研究では，例えばエルスバークのつぼの問題がよく知られている ([38])。客観的には同一の条件を与えても，合理的推論から人間の判断がずれる現象は「フレーミング効果」という名前で知られている。これらはいずれも期待値から予想される合理的判断と多くの人間の直感的判断にずれがあることを示している。

確率の直感的理解と合理的理解との乖離については，相当に数学的思考訓練を積んでいると思われる科学者でさえ例外ではない。たとえば，偏微分方程式論の分野では先駆的業績を残している数学者グランベール (1717-1783) は，2 枚の硬貨を同時に投げた場合，表が 0 枚，1 枚，2 枚出る確率が等しい，と終生信じていたといわれる ([44])。なお，彼の考え方が自明に誤りで

ある、と考えるのも誤りである。何故ならば、統計力学において、ある状況下では二つのボール（サイコロ、硬貨）を区別しないで場合の数を数える必要があるからである（ボーズ・アインシュタイン統計という。フェラー [15], 33 頁を参照されたい）。

本稿においては、前述のように、微分方程式を前提とした決定論的推論と違って、未だに様々な考察あるいは疑問が提出されている人間の直感的確率評価と、数理モデルとの乖離を取り上げて新たな説明モデルを提案するものである。人間の直感的確率評価があてにならないことは古典的確率論の大成者であり、かつ決定論の権化（ラプラスの魔という）のようにいわれるラプラス（1749-1827）が「確率の見積もりにおける錯覚について」（[25], 132 頁）、においてつとに指摘していることではあるが未だに定説をみていないように思われる。

以下本稿の内容は数理社会学会における口頭発表

(I) 確率評価の錯覚 主観的合理性の 1 側面 第 33 回数理社会学会（東京工業大学，2002 年 3 月）

および

(II) ムカデ・ゲーム再考 後ろ向き帰納法の回避モデル 第 35 回数理社会学会（慶応義塾大学，2003 年 9 月）

のレジメに加筆修正を加えたものである¹。

(I) 確率評価の錯覚－主観的合理性の 1 側面

§1.1. 動機および目的

Boudon(1989,[10],174 頁) は「主観的合理性」概念の重要性を論じる際に次のような例を引用している ([42])。

「被験者は、実験者によってコインの表－裏をあてるゲームの出る面を予測するように求められる。しかしながら、そのゲームで使われるコインは歪んでいて、表と裏は、0.8 と 0.2 の確率で出現する、と被験者は告げられる。たいていの場合、被験者は誤った解決法を選択する。すなわち、彼らは、予測されようとする一連のデータと同じ確率によって制御された出現規則を生成するのである。言い換えれば、彼らは、0.8 の確率で「表」の出現を、また 0.2 の確率で「裏」の出現を予測することを選ぶのである。そうすることによって、彼らは、0.68(註： $= 0.8 \times 0.8 + 0.2 \times 0.2$ ここでは、暗に独立性を仮定している) の確率で出現する面を正確に予測するのである。これは、もし彼らがすべての試行で表を予測することを選択する場合に得るであろう結果と比べても貧弱な結果である。なぜならば、それぞれの出現を彼らが正確に予測する確率は、この場合 0.8 と等しくなるからである。」

この事実は 1930 年代にすでに心理学の分野で動物実験、および人間について観察された現象と同一であると思われる。小野 茂 によると「確率一致現象」と呼ばれている（心理学の分野でどの程度普遍的用語かは不明 ([30], [31]。なお，[30] の 2 章「確率学習」のデータは [16] によっている。）

¹本研究の一部は、日本学術振興会科学研究費補助金（基盤研究 (B)）「秩序問題への進化ゲーム理論的アプローチ」（研究代表：大浦宏邦，課題番号：14310095，2002 年-2004 年）の援助を受けた

確率一致現象を説明する数理モデルは心理学の分野ではすでにいろいろ知られている。例えば、Estes の刺激標本モデル (stimulus-sampling model), Atkinson の強弱条件付けモデル (strong and weak conditioning model), Suppes & Atkinson の観察反応モデル (observing-response model) 等である。確率一致現象はゲーム理論の合理性と矛盾するために、「合理性」の観点から様々な議論されてきた。(H.A.Simon (1982) [37], Thrall,R.M. Coombs,C.H. & Davis,R.L. [11])

§1.2. 主観的合理性

同じ長さの線分の両端の補助線の違いによって、視覚的には同じ長さに見えない、という視覚における錯覚現象はひろく知られている。しかしながら、この事実をもって、体系的に「主観的幾何学」なるものが考察されたことはない。一方、確率、特に条件付確率に対する直感的評価が必ずしも確率の公理からの計算結果と一致しないことは古典確率論を大成したといわれるラプラス自身が強調していることである。彼のエッセイ (『確率の哲学的試論』内井惣七訳 岩波文庫青 925-1(1997)132 頁-142 頁、「確率の見積もりにおける錯覚について」) においても一章を割いて論じている。彼は「視覚に錯覚があるように精神にも錯覚がある。そして、さわってみて目の錯覚が正されるように、反省と計算によって精神の錯覚も正される」(同書 132 頁)、と主張している。

なお、Feldman ([14], 440 頁) によると「金銭でもって動機付けられると、被験者は二者択一実験で各回ごとに一番起こりそうな事象を予測することがわかった」と述べているから、一定の条件下では確率的錯覚を免れる可能性はある(丁度、視覚の錯覚が解消されるように)。ひとつの可能性として、予測することを求められているわけではなく、自己の利益を最大にすることを求められている、という違いではないだろうか。

本稿の目的は、「確率一致現象」が、数理モデルで表されるある種の「錯覚」から起こり得ることを示すことである。

§1.3. 確率評価の錯覚—数理モデルによる表現

被験者(プレイヤー I)が実験者(プレイヤー II)の提示するランダムな信号(a or b)を推量することを求められているから、プレイヤー I の利得行列は

	a	b
a	1	0
b	0	1

となる(行がプレイヤー I の戦略, 列がプレイヤー II の戦略) 実験はまず、プレイヤー I が推量して、次にプレイヤー II が信号を提示してゆくので、本来は展開形のゲームであるが、基本的なアイデアを説明するためにまず標準形のゲームとして定式化する。混合戦略を含めて表現するためにプレイヤー I の選ぶ戦略 X は確率変数とみなすことが出来る(確率概念を用いる時に分布で表現するのはよい定式化ではない)。同様にプレイヤー II の選ぶ戦略を Y で表す。ただし、被験者は実験者が次に提示する信号を推量することを求められているから、 X と Y は必ずしも独立な確率変数であるとは限らない(通常非協力ゲームでは独立性を仮定している)。つまり、 (X, Y) は $(a, a), (a, b), (b, a), (b, b)$ のどれかの値を取る確率変数である。このとき、

プレイヤー I の利得 $u_{X,Y}$ は

$$u_{X,Y} = P(X = a, Y = a) + P(X = b, Y = b)$$

である (注意: $u_{X,Y}$ は確率変数ではない.) この問題ではプレイヤー II の利得は考慮しないプレイヤー I のひとりゲームである. 従って,

問題: 確率変数 Y の周辺分布 $P(Y = a) = p, P(Y = b) = 1 - p$ を所与としたときに, プレイヤー I の利得 $u_{X,Y}$ を最大にする戦略 X を求めよ. ただし, 確率変数 X と Y は必ずしも独立とは限らない.

この問題の解は次の定理によって与えられる.

定理 1: (X, Y) の分布を次のように定めた場合に, $u_{X,Y}$ は最大値をとる.

$$P(X = a, Y = a) = p, P(X = a, Y = b) = 0,$$

$$P(X = b, Y = a) = 0, P(X = b, Y = b) = 1 - p.$$

証明: (X, Y) の分布は今の場合, パラメーター 3 つで表現されるが (根元事象が 4 個あるから) $P(Y = a) = p$ が所与であるから, 後 2 個のパラメーター (s, t) を次のように定める.

$$P(X = a/Y = a) = s, P(X = a/Y = b) = t. \quad 0 \leq s \leq 1, 0 \leq t \leq 1.$$

ここで, $P(A/B)$ は事象 B (ただし, $P(B) \neq 0$) を与えたときの事象 A の条件付確率 $P(A \cap B)/P(B)$ を表す. このとき,

$$\begin{aligned} u_{X,Y} &= P(X = a, Y = a) + P(X = b, Y = b) \\ &= P(X = a/Y = a)P(Y = a) + P(X = b/Y = b)P(Y = b) \\ &= sp + (1 - t)(1 - p) \end{aligned}$$

従って, 容易に分かるように $u_{X,Y}$ は $s = 1, t = 0$ の時に, 最大値は 1 をとる. このことは確率分布が

$$P(X = a, Y = a) = p, P(X = a, Y = b) = 0,$$

$$P(X = b, Y = a) = 0, P(X = b, Y = b) = 1 - p,$$

を意味する, つまり定理が証明された.

しかしながら, この定理は $P(X = Y) = 1$ を意味するから, ランダムに戦略を決めている相手の手を完全に予測することを意味しており, 現実には不可能である. しかし, プレイヤー I は相手の手を予測するように要請されて, 必死に予測を行なおうとする結果, 次のような錯覚に陥るであろう.

確率評価の錯覚: プレイヤー I は最良の予測と信じて次のような戦略を選択する.

$$P(X = a/Y = a) = P(Y = a), \quad P(X = b/Y = b) = P(Y = b).$$

この定式化は Boudon の次の文章を数式で表現したと見なすことも出来る ([10],175 頁)

「the subjects made the conjecture that, in order to replicate a model, a good strategy is to generate the copy by applying the very rules that govern the production of the model.」

さて、上記のような錯覚に陥った場合の論理的帰結は次の定理によって示される。

定理 2: 「確率評価の錯覚」のもとでは、確率変数 X と Y は独立である。

証明: 仮定によって、 $s = p$, $t = 1 - P(X = b/Y = b) = 1 - (1 - p) = p$ であるから、

$$\begin{aligned} P(X = a) &= P(X = a/Y = a)P(Y = a) \\ &\quad + P(X = a/Y = b)P(Y = b) \\ &= p^2 + p(1 - p) = p. \end{aligned}$$

従って、

$$\begin{aligned} P(X = a, Y = a) &= P(X = a/Y = a)P(Y = a) \\ &= p^2 = P(X = a)P(Y = a). \end{aligned}$$

故に、事象 $(X = a)$ と $(Y = a)$ は独立。残りの組合せはこの等式から自動的に独立となることが知られている（簡単な計算によって確かめられる）から、確率変数 X と Y は独立である。なお、 $P(X = a) = p$ となることから、「確率一致現象」も説明できる。本稿の数理モデルの特徴は、確率変数 X と Y の独立性がアプリアリな大前提ではなくて、モデルからの帰結として導かれることである。

確率変数 X と Y が独立となった結果、繰り返し試行においては強大数の法則が成り立ち、データの相対頻度は Y の分布と一致する。つまり、確率一致現象が観察される。ただし、繰り返し試行の数理モデルは厳密には確率変数列で表現しなければならないから、もう少し精密な定式化が必要である。

さて、もしも最初から X と Y の独立性を仮定すると

$$\begin{aligned} u_{X,Y} &= P(X = a, Y = a) + P(X = b, Y = b) \\ &= P(X = a)p + P(X = b)(1 - p) \\ &= P(X = a)(2p - 1) + 1 - p. \end{aligned}$$

となるから、 $p > 1/2$ ならば $P(X = a) = 1$ とするのが合理的選択理論においては最良の戦略である。かくて、被験者は社会学者から「君の戦略は合理的ではないね」、と言われるのである。

§1.4. コメント：問題点ならびに各種実験上の差異

(1) 確率とは何か、ということは自明なことではないと思われる。従って、被験者に確率評価を行なわせる場合、どのように問題を提示するかは重大な問題である。上記の文献に引用さ

れている実験においては、問いかげそのものの妥当性に若干の疑義があるように思われる。もともとランダムな現象を予測させることに問題がある上に、正直に予測しようとするれば、応答が提示条件と独立にはならないのが自然であろう。にもかかわらず、最初から独立性を仮定した上で最適解を「正解」と見なすのはどうもだまされた気がする。また、提示するランダム現象の確率を被験者にどう提示するかの違いが結果に本質的に関わってくる可能性は否定できないように思われる。少なくとも、3通りの違いがある。

(a) 被験者自身が客観的に確率評価できる場合。辻 ([42]) の実験においては、サイコロの目が1から4であるか、5または6であるか、を問題として提示している。

(b) 被験者に数字で確率を与える場合。Boudon に引用してある例では、確率は0.8と0.2である、と指示されている。被験者自身が状況判断したわけではない。

(c) 被験者には確率分布を教えないまま、ランダムに提示する。Feldman の例がこの場合に相当する、と考えられるが、相手が人間の場合、被験者と実験者との心理的駆け引きが行なわれる可能性がある。動物実験においては当然この状況となる。この場合、被験者は経験的に確率分布を推定せざるを得ない。動物実験においても「確率一致現象」は起こり得る、と言われている。なお、引用文献では明示的な説明がないが、どのような仕方でもランダムな列を与えたかは大きな問題である。なぜならば、これがランダムな数列だ、という定義は出来ないからである。

(2) Grant-Hake-Hornseth ([16], p.4r) では、被験者は

$$P(X = a, Y = b) = P(X = b, Y = a)$$

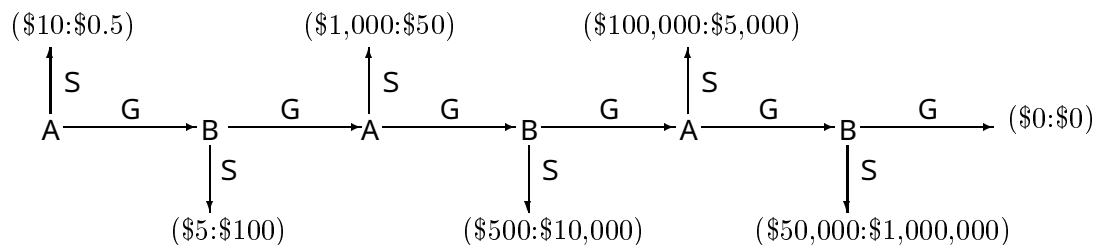
が成り立つように予測しているに違いない、というモデルを述べている。つまり、 $\{X = a\}$ でのミスの確率と $\{X = b\}$ でのミスの確率が等しくなるように推量している、というのである。独立性を仮定すれば、この等式から容易に $P(X = a) = P(Y = a)$ が導かれる。しかし、大前提として独立性を仮定しているのは問題である。

(II) ムカデ・ゲーム再考—後ろ向き帰納法の回避モデル—

§2.1. 動機

社会学あるいは心理学において、のパラドックスとかのジレンマと呼ばれている話は数理的に考えた場合、パラドックスでもジレンマでもなく正解は一意に、かつ自明に求められる場合が多い(注1)。では何故パラドックスとかジレンマと呼ばれるのであろうか。それはひとえに当該のモデルの結論が社会的常識に反し、心理的に受け入れ難いからである。社会科学における数理モデルが社会現象ないし、一般人の常識を可能な限り正確に説明しなければならない、とするならば数理モデルの前提となっている仮定を再検討し、説明力を増すようにモデルの仮定を修正する必要がある。本稿では土場 ([12]) によって論じられ、佐藤 ([35]) によって批判的に取り上げられている、いわゆる「ムカデ・ゲーム」を再考してみたい。

§2.2. ムカデ・ゲーム



ムカデ・ゲームは展開形ゲームであって、上図のように偶然手番のない完全情報2人ゲームである。先手番Aの選択枝はS（ストップ）とG（ゴー）の二つで、もしSを選択するならばゲームはそこで終了し利得ベクトル（\$10:\$0.5）を得る。左側の数字がプレイヤーAの利得、右側の数字がプレイヤーBの利得である。もしAがGを選択するとゲームは続き、Bの手番となる。Bの選択枝もAと同じSとGの二つで、意味も同じである。BがSを選択した場合、ゲームは終了し、利得ベクトル（\$5:\$100）が得られる。以下同様にゲームは終了するか続けるかが選択され、高々 $2n$ 回まで続けた後、 $2n$ ($n \geq 1$) 回目にプレイヤーBがGを選択した場合は強制的にゲームを終了して、その利得は両者ともゼロであるとする。図のとおり、ゲームが終了した場合の利得は何回目終了したかによって異なる。利得構造は次々回に分けて自分の手番にもし戻って来た時にSを選択してゲームを止めると前回の自分の手番で止めたときより利得が多い。ところが、次回の相手の手番の時に相手がSを選択してゲームが終了すると自分の利得は今Sを選択してゲームを止めた時の利得よりは低い。しかし、長くゲームを続ければ、A、Bどちらのプレイヤーにとっても初回の利得構造はパレート劣位になっているところがこのゲームのミソである。にも拘らず、最終回（ $2n$ 回目）の手番においてBは当然、合理的に判断をすると仮定すれば、Sを選択するはずであるが、その前の回（ $2n - 1$ 回目）のAにとっては、次回にBがSを選択することが明らかならば、自分の手番の時にSを選択する方が得であることは利得構造から明らかである。このように推論して行くと（逆向きの帰納法という）結局Aは初めにSを選択することが合理的である^(注2)、という結論になる。もちろん、多少なりとも相手に対する信頼ないし期待があり、相手もGを選択してくれるのではないかと、という幻想を持つのも自然な感情かもしれない（独白：しかし、それはあなたが大学行政に関わらなかった幸せな学者バカであって、このような状況で判断をする場合、最初にSを選択するのは極めて合理的であり、何の問題もない。そのように考えることの出来ない人は部局長になる資質がない、と自覚すべきであろう。なお、初めにAがSを選択することに対する現実的、心理的合理性はなにも後ろ向きの帰納法のせいだけではない。現状よりも悪いことが起こる可能性がある選択枝は絶対に選ばない、というリスク管理の観点からSを選択することはひとつの合理性である。この点も大学では日常茶飯事に経験することで、他学部相手の交渉で相手の善意をいささかでもあてにしてはいけない。しかし通常、合理的選択理論という場合は自分の利得（期待値を含む）を可能な限り最大化するように行動することが求められている。リスクを最小にする、相手との利得の合計を最大にする、相手の利得を最大にする（純粋利他行動）等を定式化する場合は、最初の利得関数を変換してから、通常の期待効用仮説に乗せる必要がある。）

この心理的葛藤は1回限りの囚人のジレンマとも共通する心理である（§4 参照）。つまり、プレイヤーの合理性と情報完備性（共有知識）の仮定のどちらか、あるいは両方に無理がある、といわれている。もっとも、「合理性」とは何を意味するのか、「共有知識」の中身は何か、といっ

たことに関しては種々議論があり、必ずしも共通了解が得られているわけではないように思われる。

このムカデ・ゲームの「パラドックス」を回避するために Aumann ([2]) が採用した着眼点はゲーム理論で通常仮定する「共有知識」を再検討することであった。「共有知識」とは、自分は相手がゲームのルールを知っていることを知っており、そのことを相手も知っており、そのことを自分も知っており、……という無限の認識列を仮定することである、とされている。社会科学における概念構成は、しばしば自己言及に陥る危険があり、ある種の宿命である、という印象がある。言葉の上での無限系列からパラドックスを導く例としてはギリシャ以来有名なゼノンの逆理^(注3)が連想される。しかし、これらの「パラドックス」からは何ら生産的結論は導かれない。その点は例えばラッセルのパラドックス^(注4)とは決定的に異なる。ラッセルのパラドックスは集合論の公理化の必要性を認識させる、という成果があった。Aumann にしろ、土場にしろ、彼らの考察から何らかの積極的な成果が得られたとは到底言い難い。

では、ムカデ・ゲームのどこを修正すれば現実感のある結論、つまり初回にGを選択することの合理性が導かれるか、ということゲーム理論の枠組みの中で考察したい。佐藤に引用がある Kreps ([24]) は合理性を緩めて、ある種の非合理性を導入することによってこの問題を解決しようとしたように思われる。このような立場は合理的選択理論の非現実性に対する批判への対案としてしばしば採用されている。しかし、Aumann([2],p.224) に指摘してあるように非合理性を定義することは困難であり、相手との協調性を仮定してしまうと最初に設定した問題そのものが消滅する。例えば、Rosenthal ([34]) はGを選択する主観確率を $\min\{1, 0.5 + D\}$ (ここで、 D は二つの選択枝の利得の差) とアドホックに仮定して、最後のプレイヤーはSを確率1で選択するが、次第に遡って、ついには最初のプレイヤーは確率1でGを選択する、という結論を導いている。仮定がかなり不自然であるし、逆向き帰納的に確率を直感的にしろ計算できるとは到底思われない。一方、共有知識の仮定(情報完備)の仮定を緩めて情報不完備ゲームによる定式化も以前から知られているが、ムカデ・ゲームの場合、情報が完備でないことから「パラドックス」が導かれているわけではない(Luce-Raiffa([26],3.6節)には「合理性」と「共有知識」の仮定が強すぎるとの批判があり、同書80-81頁にはゼロサムゲームについてムカデ・ゲームと本質的に同じ問題点を論じている。)

本稿では、形式上囚人のジレンマ・ゲームに対する Shubik([36]) の結果(割引率を持つ無限繰り返し囚人のジレンマ・ゲーム)を殆どそのままムカデ・ゲームに適用する、という定式化によって、あくまで完全合理性と(主観的、心理的ではあるが)情報完備性を仮定して、A、B両プレイヤーが初回にGを選択することが合理的である、という結論を導く(共有知識、という場合、結局は主観的、心理的であることは避けられないように思われる)

§2.3. 定式化と結論

割引率を持つ繰り返し囚人のジレンマ・ゲームの場合、よく知られているように All-D 以外にナッシュ均衡が存在してナッシュ均衡として協力が実現し得る。しかし、ある有限回 n で終わることが分かっている場合は逆向きの帰納法によって All-D 以外にナッシュ均衡はあり得ない。ここで、割引率を将来の利得を現時点で評価するから割り引いて評価する、と主観的に解釈することによって1回限りの囚人のジレンマを利得行列の違う別のゲームに変換してジレンマを回避する、という解釈がある。この定式化は、次のゲームが行われる確率が $0 < \delta < 1$ であり、その

ことはプレイヤーとは独立に決まり（例えば硬貨を投げて決める）、そのルールをプレイヤーは知っている（確率の公式から、見かけ上は無限繰り返しゲームではあるが、確率1でゲームは有限回で終了する）、という数理モデルと形式上は同じ数理構造をしている。このような定式化によって、1回限りの囚人のジレンマでも協力的行動が（主観的に）合理的である、という解釈が成り立つ。もしこの割引率が各プレイヤーごとに主観的に決まるとするならば、プレイヤーAの割引率 δ_A とプレイヤーBの割引率 δ_B は一致している必要はない。その事実を共有知識として仮定すればよい。必ずしも共有知識でなくとも、結果として δ_A, δ_B とともに $(T - R)/(T - P)$ より大であると互いに評価していれば、両者とも Trigger 戦略を取るのが合理的選択となり、結果として協力が実現する（§4 参照）。もうひとつ注意してほしいのは、この事実はある有限回 n でゲームが終了することをプレイヤーが知っている、ということとは全く異なる、ということである。無限回続くゲームは現実的ではない、という批判は「確率1で有限回で終わる」という数理モデルを誤解していると思われる^(注5)。ニュートンの力学モデルは体積ゼロで質量を持つ「質点」という仮想の概念を基本にしているが、現実の物体で体積がゼロということとはありえないと言ってニュートン力学を否定する人はいない、従って、十分長い繰り返しゲームに対して、初回においては、いつ終わるかは分からないが有限回で終わる（という「共有知識」を持っている）ゲームである、と仮定することはモデルとして不自然ではないし、その数学的帰結として現実のゲーム（十分長い繰り返しゲーム、あるいはムカデ・ゲーム）に対する現象（実験結果）が説明できるならば、そのモデルはよいモデルである、と言ってよいのではないだろうか。少なくとも「合理性」や「共有知識」の解釈をめぐる「神学論争」をするよりは生産的である、と考える（Aumann([5]) は合理性を細分して substantive rationality と material rationality に分けている。）

以上のことを頭において、ムカデ・ゲームにおいてもまず、人間は十分先のことを正確には認識できないから、ゲームはいつまで続くかはわからない、と仮定する。従って、このように仮定すると逆向きの帰納法は使えない。しかし、Gを選択した場合は利得が確定しない。Gを選択する根拠は、将来のいつかの時点でゲームが終了した時に得られるであろう期待利得が今Sを選択するより有利である、という合理的根拠に基づく確信である。実際問題としても、いつ終わるか分からないがいつかはゲームが終わるのである、ということは誰も確信しているはずである（人間の寿命のように）。ムカデ・ゲームでの問題は、初回到プレイヤーAの立場に立ったとき、次のプレイヤーBがGを選択するのが合理的と判断するであろう、とプレイヤーAが合理的に推論できるか、ということである。この点に関しては1回限りの囚人のジレンマで何故「協力」を選択することが心理的、主観的に判断して合理的か、という解釈と同じである。もしAもBもGを選択するのが合理的である、と結論されるならば3回目以降はその時点で同じことを考察すればよい。

モデルの仮定

今、ゲームを始める時点でプレイヤーAとプレイヤーBは、 $\{2, 3, \dots\}$ に値をとる確率変数 X_A と $\{3, 4, \dots\}$ に値をとる確率変数 X_B をそれぞれ主観的に心の中に持っていて、Aは X_A の分布に従って、ゲームが終了すると信じ、Bは X_B の分布に従って、ゲームが終了する、と信じていると仮定する。ゲームが終了する原因は相手の判断ミスか客観的理由によるのか、自分がその時点でやめようと思う可能性であるかは問わない。もし、この分布が共有知識となっていれば

正しく判断され、もし間違っていて評価していれば結果として不合理な選択となる。それは現実に合わせてはめる際の問題であって、モデル上の問題ではない、つまりパラドックスではない。

割引率をもつ繰り返し囚人のジレンマ・ゲームの場合も、割引率 δ は共有知識である、と暗に仮定してあるが、それぞれのプレイヤーが主観的に別々の割引率を持っていて、利得行列を変換していたとすると、一方のプレイヤーが TFT 戦略を取るのが合理的であると思って初回に C を選択しても、他方のプレイヤーは彼の主観的に合理的な判断によって初回に D を選択するかもしれない。しかし、その時点でお互いに間違っていた評価をしたことに気づくことになる。2 回目以降に割引率をどのように評価し直すのが合理的か、という基準は次の問題である。

ここで、

$$P(X_A = n) = p_n > 0, \quad \sum_{n=2}^{+\infty} p_n = 1, \quad P(X_B = n) = q_n > 0, \quad \sum_{n=3}^{+\infty} q_n = 1$$

と仮定する。分布の仮定から A にとっても B にとってもゲームは主観的にはいつかは終わる、と信じられていることになる。数理モデルは話を一般化した方がむしろ構造がよく見えて理解しやすいので、 n 回目にゲームが終了したの時の利得ベクトルを (a_n, b_n) とする。ムカデ・ゲームというときは(上記のパラドックスが生じるのは)

$$a_2 < a_1 < a_4 < a_3 < a_6 < a_5 < \cdots \quad (a_{2k} < a_{2k-1} < a_{2k+2}),$$

b_n に対しても同様に

$$b_2 < b_1 < b_4 < b_3 < b_6 < b_5 < \cdots \quad (b_{2k} < b_{2k-1} < b_{2k+2})$$

を仮定することである。このとき、A が期待効用仮説に基づく合理的判断によって初回に G を選択するための条件は

$$(*) \quad a_1 < E[a_{X_A}] = a_2 p_2 + a_3 p_3 + \cdots$$

が成り立っているときである。ここで、 $E[a_{X_A}]$ は確率変数 a_{X_A} の期待値である。同様に B が期待効用仮説に基づく合理的判断によって自分の手番の初回に G を選択するための条件は

$$(**) \quad b_2 < E[b_{X_B}] = b_3 q_3 + b_4 q_4 + \cdots$$

が成り立っているときである。かつ、この場合、相手が G を選択しているときに自分が S を選択すると平均利得が下がる、という意味で G を選択することはナッシュ均衡である。均衡解のことや後ろ向き帰納法を最初から考えないと決めて、ただ現象を説明する数理モデルを立てる、というのであれば、何時終わるかわからない無限回ゲームを想定する必要はなく、上記の期待値の計算で予め定めた n までの和と初回に S を選択した時の利得とを比較すればよい。

最初の例で、 $p_n = (1 - \delta_A) \delta_A^{n-2}$, $q_n = (1 - \delta_B) \delta_B^{n-3}$ と仮定して (δ_A, δ_B がゼロに近い程ゲームは早く終了すると予想していることになる)、実際に計算してみると

$$\delta_A > 0.00504 \implies 10 < 5(1 - \delta_A) + 1,000(1 - \delta_A)\delta_A + 500(1 - \delta_A)\delta_A^2 \cdots$$

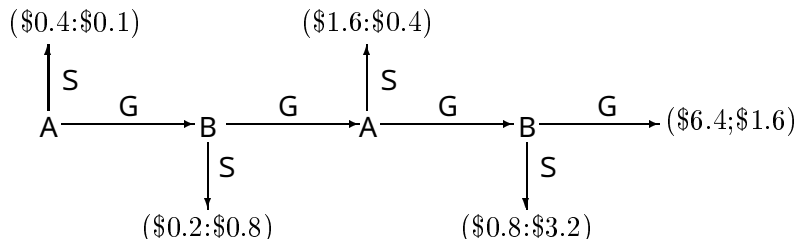
$$\delta_B > 0.00504 \implies 100 < 50(1 - \delta_B) + 10,000(1 - \delta_B)\delta_B + 5,000(1 - \delta_B)\delta_B^2 + \cdots$$

となる．ゲームがいつ確定的に終わるかわかっていない場合，または，自分ができるだけゲームを続けたいと思い，相手にもそう期待している場合は， δ_A, δ_B をあまり小さく評価するとは思えないから，上記のような仮定のもとでは二人ともGを選択するのが合理的である，と判断する可能性が高いという心理的感覚が得られる．上記の計算では…のところは無視して計算してあるから，人間の能力として無限級数を感覚的に直ちに計算できるということは仮定していない．

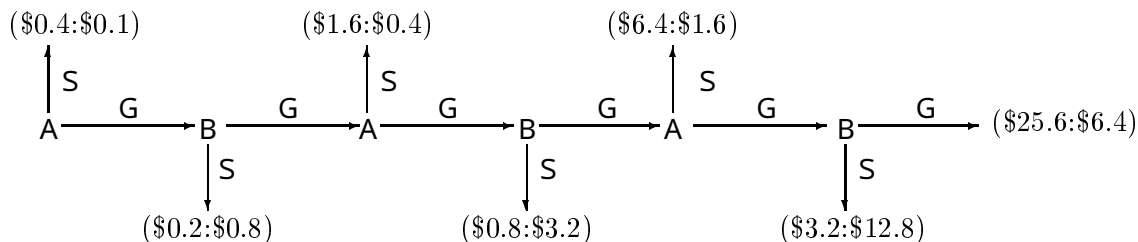
§2.4. Discussion

(1) ムカデ・ゲームの実験結果とその解析は McKelvey-Palfrey ([28]) においてなされている．彼らは，4回で終わるムカデ・ゲームと6回で終わるムカデ・ゲームを比較している．彼らのムカデ・ゲームは次のような構造をしている．

4回で終わるゲームの場合．



6回で終わるゲームの場合．



n 回目でゲームが終わった (n 回目のプレイヤーがSを選択した) 相対頻度 f_i ，および各選択枝に来た時にSを選択した条件付確率 $p_i = f_i / (1 - f_1 - \dots - f_{i-1})$ ただし， $f_0 = 0$ (原論文の定義式は間違っていると思われる．数値そのものは正しい) (ゲームの総数は 281，同じ相手とは対戦していない) は次表のようになっている (原論文 TABLE IIA, IIB より)．

4回で終わるゲームの場合．

プレイヤー (戦略)	A(S)	B(S)	A(S)	B(S)	B(G)	
相対頻度	f_1	f_2	f_3	f_4	f_5	合計
データ	0.071	0.356	0.370	0.153	0.049	0.999
条件付確率	p_1	p_2	p_3	p_4		
データ	0.071	0.383	0.646	0.754		

初回にS (ストップ) を選択した相対頻度は 7.1% であるのに対して，最終回になおGを選択したプレイヤーが 4.9% もいて，これは合理的選択理論からすれば明らかに不合理な選択である．相手が2回もGを選んでくれたことに対する義理立てなのか，自分がここでSを選ぶことの

後ろめたさなのか？あるいは、もともとゲームの意味（ルール）を理解していなかったのかも
 もしれない（被験者はアメリカ人の大学生である）

次に別のグループによる6回で終わるゲームの場合（原論文 TABLE IIA, IIB より）。

プレイヤー (戦略)	A(S)	B(S)	A(S)	B(S)	A(S)	B(S)	B(G)	
相対頻度	f_1	f_2	f_3	f_4	f_5	f_6	f_7	合計
データ	0.007	0.064	0.199	0.384	0.253	0.078	0.014	0.999
条件付確率	p_1	p_2	p_3	p_4	p_5	p_6		
データ	0.007	0.064	0.214	0.526	0.731	0.839		

6回ゲームの場合、初めにS（ストップ）を選択した相対頻度は4回ゲームより一桁小さい0.7%であり、最終回になおGを選択したプレイヤーも1.4%に減少している。ゲームの回数が増えただけにも関わらず、初めにSを選択するプレイヤーが激減している。ゲームの回数の情報は初回の選択に極めて大きな影響を与えていると推察される。最初に指示されるゲームの回数が多い場合、自分も相手もゲームを少なくともしばらくは続けたいと期待して前記の δ_A, δ_B を大きく見積もっているのではないだろうか。ゲームの回数の影響はプレイヤーBにも影響を与えていて、上記の分布を見ると6回ゲームは4回ゲームの分布をほぼ右に2回だけシフトしただけである。このことはゲームの終わりが近づくにつれてSを選択する（つまり合理的に判断する）誘引が強くなっていくことを示している。両ゲームとも回が進むにつれて選択する時点でSを選択する条件付確率は単調に増加していることが見て取れる。とは言っても最後の選択肢でもなおGを選択するプレイヤーは少数ながら存在する。この現象を説明するためには何か別の動機ないし心理的要因を仮定する必要があるだろう。たとえば、人間は「自分が利己的な人間だと思われたくない、そう思われると不利である」という利己主義の心理が働くのかもしれない。

彼ら自身も数理モデルを立ててこの現象を説明しようと試みている。彼らの定式化は 1. 情報不完備ゲーム（人間のタイプとして利己主義者と利他主義者がいる）2. 2種類のエラー（行動のエラー、信念のエラー）を仮定する、というものである。従って、本稿とはモデルの設定が基本的に異なる。

(2) 囚人のジレンマゲームとの対比：1回限りの囚人のジレンマに対する実験的研究とその理論付けに関しては山岸グループによる精力的研究がある（[18],[21],[40],[43],[45]）。[45]によれば、1回限りの囚人のジレンマ・ゲームにおいても協利行動をとるのは「社会的交換ヒューリスティクスが活性化されることによって、ゲーム自体が安心ゲームの利得構造に変換されて認知されるためである」という説明がなされている。なお、山岸達は「繰り返しゲームとは特定の相手と繰り返し行うゲームである」、と定義しているが、相手が違うにしろ複数回のゲームの利得の合計を考慮する限りは繰り返しゲームであると考えの方が自然であると思われる。例えば、有名な割引率を持つ繰り返しゲームは、現実には1回限りのゲームでも人間が社会的動物である以上、常に将来再びであうかもしれない、同じ相手と再びプレーする確率は δ くらいであろう、と心の中で想定した利得を基に意思決定している、と解釈することも可能であろう。その場合、例えば、TFT 戦略対 All-D 戦略との利得構造は安心ゲームと同様の利得構造になる。違いは山岸達がゲーム理論に社会心理学的根拠を導入しているのに対して、割引率を持つ囚人のジレンマ・ゲームはあくまでゲーム理論の枠内のみで「ジレンマ」を解決しようとしていることである。ゲー

ム理論の枠の外のからもっともらしい説明をすることは他にも可能であろう。例えば、[13],[27]のように感情自体が例えば脳内物質によって引き起こされて、進化的淘汰の対象になり得ると仮定すると囚人のジレンマ・ゲームにおける協力行動を説明する別の解釈はあり得る。すなわち、[45]に述べられているように囚人のジレンマ・ゲームにおいて、相手が先にすでにCを出していることが分かった後でもなお、Cを出すプレイヤーが実験的に多数見られることの説明は「Dを出すことに対する罪悪感があるから」であり、「罪悪感」をある程度持つグループは協力行動をとりやすく結果的に個人的にもグループ的にも進化的に有利である、という解釈である。この解釈を実証するためには罪悪感を引き起こす脳内物質を特定する必要がある、現在の研究レベルをもってしても容易には検証できないかもしれないが^(注6)。もちろん、罪悪感が強すぎると生物学的な適応価を下げるであろうから、その場合は「宗教」ないし「神」を「発明」して「救い」や「赦し」の概念を導入する必要がある。[1]には、大学生292名に対する自記式調査から、罪悪感には社会的適応機能がある、との結果が報告されている。

なお、[17]の題名は One-Shot PD の協力行動となっているが割引率を持つ繰り返しゲームの利得行列を用いている。ということは割引率を「実際に」ゲームを繰り返す確率、という解釈をしていない、ということであろう！「1回限りの囚人のジレンマ・ゲーム」とは何を意味するか、という定義からして問題であることがわかる。数理モデルが表現している「1回限りの囚人のジレンマ」以上に何がしか心理的、社会的プラス・アルファの要素を持ち込むならば、という条件の下での囚人のジレンマ・ゲームの研究、と明示すべきであろう。

(3) 1回限りのゲームではなく、予め定められた有限回の囚人のジレンマ・ゲームの場合も(1回でも協力行動を選択するプレイヤーが多いわけであるから)当然初回にCを選択するプレイヤーはより多くなる。しかし、合理性と完全情報を仮定する限り、後ろ向き帰納法によって、All-Dが唯一のナッシュ均衡解である。それでもなお、協力行動をとることの合理的説明もいろいろなされているが([6])、ゲーム理論の枠内で有限回のゲームのままの説明にはやはり無理を感じる。ゲームの回数が十分長い場合は、割引率(次のゲームが行われる確率)を持つ無限回ゲーム(確率1で有限回でおわるが、回数を予め指定することは出来ない)という数理モデルによって、協力行動が発生する、という説明の方が数理モデルとしては自然な感じを持つ。

なお、従来の論文では有限回あるいは無限回繰り返しゲーム、という言葉の意味にかなりの混乱が見られる。数理モデル上は a) 予め指定された回数ゲームを行う、b) 確定的に、ゲームが終了する回数を指定することは出来ないが、確率1で有限回で終わるゲーム、c) 正の確率で(あるいは可能性として)無限回ゲームを行う(この場合の利得を確定するためには極限操作が必要)の違いは明白である。多くの論文では b) と c) を共に無限回ゲームと呼んでいるようであるが、ゲームを行う回数が有限である、という意味では b) は有限回ゲームであるとみなすべきであろう。たとえば、有名な割引率を持つ繰り返し囚人のジレンマ・ゲームは、頭の中でゲームが無限回続くと想像して将来の利得を割り引いて現在のゲームの利得構造を変更する場合は空想によって利得行列を変更するだけであるから1回きりのゲームとみなされるが、もし次回にもゲームが行われる確率が $0 < \delta < 1$ である、と理解した場合は b) のケースなのである。しかしながら、多くの論者は c) のケースと理解しているようである。これは可算加法性を前提とするコルモゴロフの確率論に立つ限り誤解である。かといって素朴なラプラス流の有限確率論では数学的に正確に記述することはできない。純粹数学の分野以外ではコルモゴロフの公理的確率論以外の「確率」が種々議論されているが現実理解に適用してみても十分に成功しているとは思われ

ない。なお、不確実さを確率論的に定式化するのではなく、いわゆるファジー理論によって理解しようとする試みはあるが本稿では取り上げない。

(注1) Shubik([36],p.193)の最後のフレーズには

The paradox of the Prisoner's Dilemma will never be solved—or has already been solved—because it does not exist. とある。

(注2) Phillip-Sugden([32])は n 回繰り返し囚人のジレンマゲームにおいて、後ろ向き帰納法を用いた結果として初回にDを選択することが必ずしも合理的とは言えない、ということを主張している。

(注3) 先にいるのろまな亀に足の速いアキレスは永久に追いつくことが出来ない、というパラドックス。

(注4) すべての集合を集めたものを集合として扱うと矛盾が生じることを指摘した。

(注5) Basu ([6]), Rosenthal([33], Neyman([29])も長さが有限の繰り返しゲームを扱っているが本稿とは定式化が全く異なる。

(注6) 「不安」を引き起こす脳内物質は知られている([20])。「不安」「驚き」「怒り」「喜び」「嫌悪」「悲しみ」は人類共通の感情として脳の機能から説明する説はある([39])。残念ながら「罪悪感」はこの中に含まれていない。しかし「羞恥心」も抜けているのではないだろうか(「羞恥心」「罪悪感」はバイブルの記述から自明なのかもしれない)

謝辞：本稿は広島修道大学経済学研究第8巻第2号1 19頁(2005)より転載の許可を頂いたものである。

参考文献

- [1] 有光興記(2001) 罪悪感, 羞恥心と性格特性の関係 『性格心理学研究』第9巻 第2号 71-86.
- [2] Aumann,R.J. (1992) "Irrationality in Game Theory". in Dasgupta,P., D.Gale, O.Hart, and E.Maskin.(eds.) *Economic Annals of Markets and Games: Essays in Honor of Frank Hahn*. Cambridge, Mass.:MIT Press.
- [3] ———. (1995) Backward induction and common knowledge of rationality. *Games and Economic Behavior*, vol.8, 6-19.
- [4] ———. (1996) Reply to Binmore. *Games and Economic Behavior*, vol.17, 138-146.
- [5] ———. (1998) On the centipede game. *Games and Economic Behavior*, vol.23, 97-105.
- [6] Basu,K. (1987) Modeling finitely-repeated games with uncertain termination. *Economics Letters* 23,147-151.
- [7] ———. (1990) On the non-existence of rationality definition for extensive games. *International Journal of Game Theory*, vol.19.pp.33-44.

- [8] Binmore,K.(1987) Modeling rational players. *Economics and Philosophy*, vol.3, 179-214.
- [9] ———. (1996) A note on backward induction. *Games and Economic Behavior*, vol.17, 135-137.
- [10] Boudon,R.(1989); Subjective rationality and the explanation of social behavior. *Rationality and Society*, vol.1,No.2,pp.173-196.
- [11] Thrall,R.M. Coombs,C.H. & Davis,R.L.(Eds.)(1954); *Decision Processes*. Wiley & Sons.
- [12] 土場 学 (1993) 「ゲームのパズルー合理性のパラドックスと知識のパラドックス」海野道郎編 『社会的ジレンマに関する数理社会学的研究』科学研究費報告書 .
- [13] 遠藤利彦 (1996) 『喜怒哀楽の機嫌 情動の進化論・文化論』岩波科学ライブラリー 41 . 岩波書店
- [14] Feigenbaum,A./Feldman,J.(1963); *Computers and Thought*. McGraw-Hill. 『コンピューターと思考』安部統, 横山保監修, 好学社 (1969), 第II部第3節「2者択一実験における行動のシミュレーション」(J.Feldman)pp.425-450.
- [15] Feller,W. (1957) *An Introduction to Probability Theory and its Applications*, vol.1. 『確率論とその応用』I・上 河田龍夫監訳 現代経営科学全集5 紀伊国屋 (1960)
- [16] Grant,D.A. Hake,H.W. & Hornseth,J.P.(1951); "Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J.Experimental Psychology*, vol.42,no.1,1-5.
- [17] Harrington,Jr.J. (1995) Cooperation in a one-shot Prisoner's Dilemma. *Games and Economic Behavior*, vol.8, 364-377.
- [18] 林 直保子 (1995) 繰り返しのない囚人のジレンマの解決と信頼感の役割 *The Japanese J. of Psychology*, Vol.66, No.3, 184-190.
- [19] Humphreys,L.G.(1939); The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J.Exp.Psychol.*,vol.25,pp.141-158.
- [20] 貝谷久宣 (1997) 『脳内不安物質』ブルーバックス B1184 講談社 .
- [21] 清成透子, 山岸俊男 (1996) コミットメント形成による部外者に対する信頼の低下 *The Japanese J. of Experimental Social Psychology*, Vol.36, No.1, 56-67.
- [22] Kolmogoroff,A.N.(1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer 『確率論の基礎概念』根本伸司他訳 東京図書 (1969)
- [23] 河野 敬雄 (2001) : 条件付確率に関する2, 3の注意 「数学」53巻4号, pp.90-98. 岩波書店
- [24] Kreps,D.M. (1990) *Game Theory and Economic Modelling. Chapter 4*. Oxford University Press.

- [25] Laplace,P-S.(1814); Essai philosophique sur les probabilités. 『確率の哲学的試論』内井惣七訳 岩波文庫青 925-1(1997). 132 頁-142 頁 , 確率の見積もりにおける錯覚について
- [26] Luce,R.D. and Raiffa, H. (1957) *Games and Decisions, Introduction and Critical Survey*. pp.80-81. Dover Pub.
- [27] マット・リドレー (2000) 『徳の機嫌 他人をおもいやる遺伝子』岸 由二監修 翔泳社 .
- [28] McKelvey,R.D. & Palefrey,T.R. (1992). "An experimental study of the centipede game. *Econometrica*, Vol.60,No.4, 803-836.
- [29] Neyman,A. (1985) Bounded complexity justifies cooperation in the finitely repeated Prisoner's Dilemma. *Economics Letters*. Vol.19, 227-229.
- [30] 小野 茂 (1966); 『学習実験』, 情報科学講座 E.17.3. 北川敏男編 , 共立出版
- [31] ——(1976); 『心理学における数学的方法』, 「現代の心理学 4」培風館
- [32] Phillip,P. & Sugden,R. (1989) The backward induction paradox. *The Journal of Philosophy*. Vol.LXXXVI, No.4,169-182.
- [33] Rosenthal,R.W. (1980) New equilibria for noncooperative two-person games. *J. of Mathematical Sociology*. Vlo.7,15-26.
- [34] ———. (1981) Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, vol.25,92-100.
- [35] 佐藤嘉倫 (1998) 「合理的選択理論批判の論理構造とその問題点」『社会学評論』49 卷 2 号,18-35.
- [36] Shubik,M. (1970) Game theory, behavior, and the paradox of the Prisoner's Dilemma: three solutions. *Conflict Resolution*. vol.XIV, No.2, 183-193.
- [37] Simon,H.A.(1982) *Models of Bounded Rationality*, vol.1,2. MIT Press.
- [38] 繁榎算男 (1995) 『意思決定の認知統計学』朝倉書店
- [39] スーザン・グリーンフィールド (2000) 『脳の探求』新井康充 , 中野恵津子訳 無名舎 (2001) .
- [40] 高橋伸幸 , 山岸俊男 (1996) 利他的行動の社会関係的基盤 1996 *The Japanese J. of Experimental Social Psychology*, Vol.36, No.1, 1-11.
- [41] Todhunter,I. (1865) "A History of the Mathematical Theory of Probability - from the time of Pascal to that of Laplace-" Macmillan co. 『確率論史-パスカルからラプラスの時代までの数学史の一断面-』, 安藤洋美訳 現代数学社 (1975)
- [42] 辻 竜平 (1991); 「常識的思考に関する研究—主観的合理性の概念をふまえて—」関西学院大学社会学部卒業論文.

- [43] 渡部 幹, 寺井 滋, 林直保子, 山岸俊男 (1996) 互酬性の期待にもとづく1回限りの囚人のジレンマにおける協力行動 The Japanese J. of Experimental Social Psychology, Vol.36, No.2, 183-196.
- [44] 山崎英三 (1972) ダランベールと確率論 (I)(II) 科学史研究第II期 11巻 No.101, pp.17-24, 同 No.102, pp.65-72.
- [45] 山岸俊男, 清成透子, 谷田林士 (2002): 社会的交換と互惠性 なぜ人は1回限りの囚人のジレンマで協力するか 『進化ゲームとその展開』10章 佐伯 胖・亀田達也編著 共立出版