HOKKAIDO UNIVERSITY

Title	集団内協力と評判心理
Author(s)	山岸, 俊男
Citation	新世代法政策学研究, 10, 109-144
Issue Date	2011-02
Doc URL	https://hdl.handle.net/2115/45067
Type	departmental bulletin paper
File Information	HJNGLP010_007.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

集団内協力と評判心理

山　岸　俊　男

心の本質的社会性

（山岸） どうもご紹介ありがとうございます。山岸でございます。前回いつでしたっけ、しばらく前に、法学部の方でお話しさせていただきまして、今回またCOEの方ということで、少し内容を変えないといけない。ということで、ちょっと苦労したんですけれども、今日お話しさせていただきますのは、最初にまず、文学部の方のCOEがとりあげています、「心の社会性」というテーマなんです。

心の社会性というのは、いったいどういうことを意味しているのかということを、話の前半に概要をお話しさせていただきます。私がやっています信頼の研究は多少ご存じの方もいらっしゃるかと思うんですけれども、実は信頼の研究は10年ぐらい前で、もうほとんどやめたような状態でして、実は最近もう少し別のといいますか、ちょっと広がった研究を中心にやっています。それが集団の話とか利他性の話とか互恵性の話とか、そういう話なんですね。話の後半では、そちらの方でいったいどういうことをやっているのかというのを、いくつか実験の例を紹介しながらお話しさせていただきます。

ですから、皆様方が私に期待されている話とちょっと違うかもしれないんですけれども、お許しください。また、こちらのCOEの研究とどうつながるかは、ちょっと分かりませんけれど、あまりつながらないかもしれません。そうしたらごめんなさい、申し訳ありません。

私どものCOEは「心の社会性」というテーマで行っております。それは基本的にどういう問題かといいますと、心の社会性の進化、適応的基盤

を明らかにするということですね。人間の心、「人間」というのは、通常、我々が社会科学で扱っている人間ですね。文化的あるいは社会的存在としての人間です。これを片仮名で「ヒト」と書きますと、これは生物学的な意味での生物としての人間を意味します。

普通は人間、こちらの漢字を使った方は、本質的に社会的存在であると考えられていますけれども、実は生物としての、つまり生物種としてのヒトも本質的に社会的なのであるという基本的な発想が、ここでいう心の本質的社会性にあたります。それがなぜそうなのかとか、どのようにそうなのかということを研究したいというのが、我々の文学研究科の方のCOEの基本的な狙いなんですね。そのことが、つまり生物種としてのヒトの本質的な社会性というものが、私たちの社会の在り方にどのようにかかわっているんだろう、そういうことを広い意味で考えています。

心の本質社会性ということを言いますと、それは古くからの問題じゃないかと、何が新しいんだよ、そんなの、と思われるかもしれません。これは実は古くて新しい問題だと私どもは考えています。古くてというのは、これは生物種としてのヒトは文化によって人間になるという意味での社会性です。これは現在ある社会科学の１つのパラダイムですね。

人間の心は生まれたときには白紙の状態であって、そこに文化が注入されることで生物としてのヒトが、他の動物とは区別される人間となるんだと、こういう一般的な理解が広く20世紀後半の社会科学に受け入れられてきました。タブララサの人間は生まれたときには白紙の心を持っている。人間の心はホワイトボードだというわけですね。ホワイトボードにいろいろと書き込まれて人間に育っていくんだと。

この考え方は、実は20世紀の社会科学の大きな足かせになってきたのではないかと私たちは考えています。それは、この白紙の心という考え方は、言い換えれば生物種としてのヒトに特有の人間性なるものはないということを意味しているからです。

このことは言い換えれば、生物種としての人間性について考えることは、まったく意味がないのであるということを意味しています。だから人間性についての研究なんてどうだっていいんだよ。それは生物学者が研究するかもしれない。だけどそんなことは、社会を研究する社会科学者にとってはどうでもいいことであるという、一般常識みたいなものが広く受け入れられてきたと思うんです。

その代表選手がよくご存じのマーガレット・ミードですね。切手になったりして、すごく有名な方です。たぶん高校の教科書にも出てくる人ですね。彼女が書いた本に、『Coming of Age in Samoa（サモアの思春期）』というすごく有名な本があります。

この本でミードが何を主張したかというと、サモアの社会には競争も抑圧もない。性が開放されているために青年期のストレスや思春期のストレスや葛藤がない。つまり荒れる子供たちはサモアにはいないんだと。競争がなくて、よくできる者は仲間が追い付くまで待っているといったことも書いてある。

争いが少なく戦争もほとんどない。個人に重大な反抗期を起こすほどの抑圧がないと。だから事実上自殺も存在しない。子供への叱責やせっかんはまれであると。性交を娯楽と見なしているので、特定な相手に対する執着心も嫉妬心もない。結婚前の青少年は自由な性行為を行うと、だから強姦も存在しないと、こういうことをおっしゃいます。

このインプリケーションは、だからご覧なさいと、西欧の社会で当たり前だと思われていることが当たり前じゃない社会があるんですよ。だから文化次第で人間性なんてものはいくらでも変わるんだと。つまり人間性なるものは、文化によってつくられるものであって、従ってどのような人間性もありなんだよということを言うわけです。

このミード的な考え方は、だいたい20世紀の中ごろに一般に広まって、それ以降20世紀後半の社会科学で広く受け入れられています。たぶん多くの社会科学の研究は、これを直接引用しているわけじゃないんですけれども、一種の常識として、そんなことは決まっているでしょうというぐらい広く受け入れられてきたわけですね。

そのことが持っているインプリケーションは何かというと、人間性についての科学的理解を無視した形で社会科学が発展してきた、ということです。逆に言えば、人間性を無視する形で社会科学がつくられてきたために、社会科学が自然科学に圧倒的な差をつけられてしまった。20世紀の後半に自然科学は大きな発展を遂げたんですけれども、社会科学は停滞してしまった。それはなぜか、それは人間性についての科学的理解を無視してきたからだというのが私の１つの解釈なんですね。

面白いことに、実はこのミードの研究に対しては、デレク・フリーマンという人が最近になって緻密な研究をしています。実はマーガレット・ミードはサモアに６カ月しかいませんでした。しかも現地の人々の間では暮らしていませんでした。どうしてかというと、そんな汚いところに住むのは嫌だからと言って、アメリカ人の医者だったか宣教師だったか忘れちゃいましたけど、そこの家に住んで、そこからあんまり出ないで、そこにいろいろなインフォーマントに来てもらって話をしてもらったわけですね。

そういうインフォーマントの話を基に『サモアの思春期』という本を書いたんですけれども、このデレク・フリーマンという人が、もちろんほかの人もたくさんやっていますけれども、ちゃんと当時のサモアの社会の状態を調べてみると、それはいろいろな記録がありますからいろんなことが分かるわけですよね。

例えば戦争がないといいますが、実際にミードがサモアにいたときに戦争が起こっています。それからサモアは位階社会というハイアラーキカルにオーガナイズされた社会であって、誰もが上位に上りたがっている、と。上位の者に対する反抗も許されていません。当然、自殺も存在します。子供に対してひどいせっかんもしている、と。処女であるということは非常に重視されていて性交を娯楽と見なすようなことはあり得ないんだ、等々、いろいろと反論しているわけですね。

これは本当のことかどうか分かりませんけれども、実はミードに話をしたという、今、老人になっちゃっていますけれども、そういう人がいて、あれは冗談で言ったんだよということを言っているんです。それは本当のことかどうかは私には分かりません。だけどこういう違いを見ると大いにあり得ると思うんですよね。これが１つの今の状態を象徴していると思います。

さきほど、心の社会性というのは古くて新しい問題だと言いましたけれども、その新しいということはどういうことなのか？新しいということは、生物種としてのヒトの心は、ヒトの体と同様に進化によって作られた適応のための道具であるという考え方ですね。もちろん脳が進化の結果作られたということは、これは誰でも認めます。しかし、人間の脳の中で起こる情報処理ですね、これが心なんですけれども、脳が進化したのは心がある役割を果たさないといけない状態があったからこそ進化したわけです。

人間の脳は、これはすごく高価な臓器なんです。これを体重比で言うと、ほんの数パーセントの重さなんですけれども、エネルギー消費量が２割ぐらいあるんです。だから一生懸命本を読むとダイエットにいいんじゃないかなと思うんですけれども、それほど維持費がかかる臓器なんですね。非常にエクスペンシブな臓器を人間が進化させた以上は、それに見合うような役割を果たしてきたはずなんですね。それっていったい何なんだろう？

脳があるだけではまったく意味がないので、脳があることによって心が働くということが大事です。その心の働きこそが脳の進化を生み出してきた。だから心の働きそのものも進化の結果であるということになります。だから人の心には、人間性という、進化の過程によってつくり出された人間性が埋め込まれているはずなんです。つまり人間の心は人間の体と同様に、人間が置かれた、人類が置かれた淘汰圧の下で進化した適応のための道具なんですね。このことは何を意味するかといいますと、心の可塑性には限界があるということです。

例えばはさみとのこぎり、これはそれぞれ道具なんですけれども、はさみにできることと、のこぎりを使ってできることは違いますよね。大きな木を切ったりするためには、はさみではどうしようもありません。だからのこぎりを使わないといけない。だけど、のこぎりではできない細かな作業にははさみを使う。

モジュール脳という考え方があります。人間の脳は、人間が繰り返し直面してきた適応問題を解決するために特化されたモジュールから構成されているという考え方ですね。モジュールというのはマイクロチップと考えてもいいんだと思うんですけど、人間の脳は巨大な電子計算機、コンピューターなんじゃなくて、マイクロチップがいろいろと埋め込まれている。それぞれのマイクロチップは、それぞれ人間が繰り返し直面してきたそれぞれの適応問題に対応しているという考え方ですね。

だからはさみに対応するようなモジュールがある。のこぎりに対応するようなモジュールがある。例えば社会関係を処理するためのモジュール、あるいは異性を獲得するためのモジュールとかいろいろとあるわけですね。文化差って何だといいますと、もちろんはさみにもいろいろなはさみがあります。同じはさみの役割をしながら、西洋ばさみと和ばさみでは、これは形が違いますよね、だけどどちらでもその目的は達成される。そう

いった形での心の可塑性というのは当然あります。

ところが人間がのこぎりを使って適応しないといけない問題に繰り返し直面していったときには、それに適したモジュールが進化したはずであると、そう考えますね。これが進化心理学の基本的な発想なんです。

人の心の社会性って何なんだろうということを考えるときに重要なのは、人間が適応しないといけなかった淘汰圧ですね。人間にとっての淘汰圧は、基本的には自然環境よりも社会環境なんです。人間にとって一番重要な適応環境は社会環境であったということですね。面白い研究があって、いろいろな類人猿を比べて、脳の大きさと、それから集団の大きさを比較すると、集団の大きさと脳の大きさには非常に強い関係があるんです。

だから脳が大きくなっているということは、複雑な社会関係をうまく処理する必要があったから脳が大きくなったと考えられています。これは社会脳という基本的な発想なんですが、マキャベリアンインテリジェンスという言い方もしますけれども、そういう意味で、人の心は主として社会環境への適応の道具なんだと考えられます。

ヒトの脳の進化を考える場合には、道具をうまく使えるということも重要なんですけれども、それともっと重要なのは、ちゃんと人間関係を把握しておいて、だまされないようにしながら、ほかの人間と協力しないといけないということがとても重要なんです。一方的にお人よしになってはだまされてしまう。利用されてしまう。そうすると、そんな人間の子孫は残っていきませんから、結局だまされないようにしながら、うまく他人と協力して社会をつくっていくような心の仕組みが進化しているはずであると。それが我々が考えています人の心の本質的な社会性であるということになります。あ、言い忘れていましたが、話している間に質問があったらいつでも質問してください。

これが私どもが主張している新しい意味での心の社会性なんですね。だから心の社会性というのは古くて新しいものなのです。古い心の社会性は文化決定論、だけど新しい意味での社会性は、人間の心は社会環境への適応のために進化してきた情報処理の仕組みである、そういう考え方です。

互酬性（互恵性）と利他性

そういう心の社会性を考えるときに、たぶん一番重要なモジュールといいますか心の働きは、互恵性ないし互酬性と利他性であろうと思っています。そうじゃないという主張をされる方もいますけれども、私はそう思っています。ということで、今日は残りの時間を使わせていただいて、人間の心の社会性の中でも、特に互酬性と利他性を中心にいろいろな実験の話を交えながら話をさせていただきます。

まず最初に、互恵性と利他性の中の互恵性の話なんですけど、互恵性あるいは互酬性というと結局お返しをするとか、お返しを期待して行動するということですね。そんなものは当たり前じゃないかと思いますけれども、実はそんなに当たり前のことではないんですね。

人間以外の動物でこういったお返しをする、あるいはお返しを期待して相手に利益を供与する、そういう動物はいるんだろうかというと、おとぎ話とか童話の中ではいますね。鶴の恩返しとかいろいろと動物がお返しするといわれていますけれど、本当にこういったことをするかというと、これは実は、今、霊長類研究者の中ですごくホットなイッシューですね。すごく面白い研究がいっぱいあります。

今まで人間以外の動物の互酬行動の例としてどんなものが挙げられているか。まず、数年前までよく引用されていたのがコウモリなんです。コウモリというのは、僕は生物学者じゃありませんから詳しいことはよく知らないんですけど、あれはどうもいつも血を吸っていないといけないんですね。１日、２日、血を吸えないと死んじゃうんです。かわいそうな暮らしだと思うんですけれども、血を吸えなかったコウモリがすぐ死んじゃうのかというとそうではなくて、ほかのコウモリから血を分けてもらいます。たくさん吸ってきたやつがいて、ちょっと分けてやる。

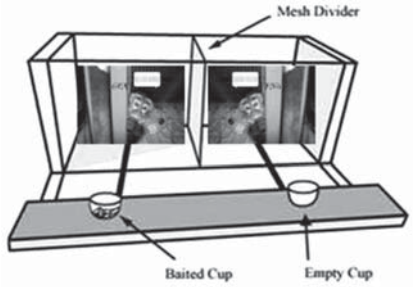
そうすると、そのことを覚えていて、ちゃんとお返しをするんだよ、という研究がよく知られていました。ところがどうも最近の研究では、それは違うよということになっているみたいなんですね。結局それは、血縁関係にあるコウモリが、お互いに一方的に利他行動をするんだということのようです。利他行動の説明は、１つは血縁淘汰の話なんですね。これは例えばリチャード・ドーキンスの『利己的な遺伝子』という本が、しばらく

の間すごく有名になりましたのでご存知の方も多いと思いますけど、その考え方です。

基本的に血縁でつながっている他個体を助けるということは、自分自身の遺伝子を助けるのと同じことである。だからそういう血縁関係にある他個体を助けるような行動を取らせる遺伝子を持っていると、その遺伝子自身が助けられるので、そういった遺伝子が拡散していく。だからそういう行動が進化するんだという考え方です。結局、血吸いコウモリの場合はお返しじゃなくて、そういう血縁淘汰によって説明されるような行動なんだということになっているようです。

それからよく使われるのが猿の毛繕いなんですね。毛繕いをしてやると毛繕いをしてもらえる、と。実際にそういうことがあるんですけれども、これを互恵行動、互酬行動と呼ぶかどうかは、私自身は非常に問題だと思っているんですね。単なるミューチュアリズムかもしれない。

つまり、毛繕いをすると相手が自分を攻撃しなくなる、だから毛繕いをしてやる。毛繕いをしてもらうと気持ちがいいものですから、それに対する敵意が減るので、そうすると毛繕いをする。単にみんながそうしているだけじゃないかという考え方もあります。ここはちょっと考え方が難しいところだと思うんですね。



De Waal and Berger, *Science*, 2000

それから最近は、霊長類学者の人たちがよく実験をします。非常に面白い実験がいっぱいあるんですね。一番よく知られている実験の１つが、カプチンモンキーをつかった実験。カプチンモンキーというのは日本語にするとフサオマキザルという猿なんです。このカプチンは社会性がすごく高いことで知られています。チンパンジーなんかよりずっと社会的な猿なんです。

この猿を使って、こうしたおりの中に２匹の猿を入れておきます。ここは金網で仕切られています。こちらにこういう板の上におわんを置いておいて、おわんの中に、こちらのおわん（baited cup）には食べ物を入れて

います。こちらのおわん（empty cup）は空です。左側のお猿さんは、この食べ物を食べたいので、これを一生懸命引っ張るわけですね。だけど1匹が引っ張っただけでは板は動かないようにできている。２匹が一緒に引っ張らないと餌が近づいてこない。よく考えた実験だと思うんですよね。

そうすると、こちらのお猿さんが助けてくれるか？自分は餌をもらえないんですよ。餌をもらえないのに、こちらが一生懸命引っ張ろうとすると助けてくれるか。この実験はドゥ・ヴァール、最近『共感の時代へ』という新しい本が日本語に訳されたりしていて、それ以外にも翻訳が多い有名な方なんですけれど、このドゥ・ヴァールさんという方が行った実験です。ドゥ・ヴァールさんという方は、ともかく動物の利他性を強調する方で、霊長類には利他性があるんだということを一生懸命言っている人ですが、彼らがやったこの実験では、右側のサルは左側のサルを助けてくれるんです。

一番面白いのは、助けたときに、左側のお猿さんが助けてくれたお猿さんに餌を分けてやるかどうか。分けてあげるのが互恵行為ですね。助けてくれたお礼に餌を分けてあげるかどうかですから。彼らは、この実験では、助けてもらうと餌を分けるんだと言っています。実験結果は、そう解釈できる結果なんですね。

ところが、よく実験結果を見てみると、「はい、ありがとうございました。これがあなたの分ですよ」と、こう分けるわけじゃないんです。どうするかというと、右側のお猿さんが、左側のお猿さんが取って来た餌に手を出して取っていこうとするんですね。それを積極的に妨げることをしないということ。だからこれは解釈が結構難しいと思うんですけれども、これがお猿さんの恩返しの一番有名な実験例です。だから、お猿さんに互恵性があるかどうかを実験で調べるのは難しいですね。

後でもう少しいろいろな実験を紹介しますけど、お猿さん、あるいはチンパンジーは、積極的に助けるということはほとんどないんです。親が子供に対しても積極的に渡すということもありません。食べ物を子供には積極的に渡さないんですよね。子供が取っていくのを邪魔しないということなんです。

それは助けているということになるんだろうなとは思うんですけれども。次に紹介するのは、霊長類研究所の山本さんというポスドクの人がや

っていた実験です。最近、霊長研の岡山にある林原という研究施設に、ボノボの研究センターができました。今までチンパンジーの研究が中心だったんですけれども、やっと日本にもボノボの研究ができる場所ができました。

チンパンジーとボノボというのは、外から見るとほとんど見分けがつかないぐらいよく似ているんですね。人間の先祖がチンパンジーの先祖と分かれたのが500〜600万年ぐらい前なんだけど、それ以後にボノボとチンパンジーが分かれたので、だからかなり近いんです。だけどボノボとチンパンジーというのは社会性に関してかなり違うんですね。チンパンジーは基本的にハイアラーキカルでボスが支配する。それに対してボノボは基本的に平等なんです。

セックスをめぐっても基本的に平等です。女の方が積極的だというか、いろいろなコンフリクトが起こりそうになると、セックスをして争いを避けるということをします。一昔前のヒッピーの時代のスローガンみたいなことをやっているんですよね。山本さんは、今はボノボの研究センターの助教になっていますけれど、彼は犬山の霊長類研究所で面白い研究をやっています。

この実験では、まずチンパンジーを訓練して自動販売機を使えるようにするんです。自動販売機の穴にお金を入れると餌が出てくる。だたし、餌は隣のおりに出てくるんです。まず自分の部屋の自動販売機にお金を入れて、隣の部屋に行って餌を食べるという訓練をします。みんなそうできるようにするんですね。

その後で、二つの部屋を仕切っちゃって移動できないようにして、それぞれの部屋に同じような自動販売機を入れておく。そして、それぞれの部屋に１匹、１匹と言うと霊長類研究所の所長の松沢さんに怒られるんで、１人と言うことにします、チンパンジーは人間と同じだからもう１人と言わないといけない、こちらにももう１人チンパンジーを入れておいて、お互いの部屋は仕切っちゃっていますから、相手のおりに出た餌は自分では食べられません。お互いに相手がコインを入れてくれると自分が餌を食べられるんですね。

このときにお互いにどうするかを調べるのが実験の目的です。たぶんこれが人間だったら、たちまち、お互いにコインを入れあうようになりますよね。ただし、子どもたちが何歳ぐらいになるとお互いにコインを入れあうようになるのかは、まだわかっていません。そうした実験をやってみると面白いとは思うんですね。それに類似した実験は私もやっていますけれども。

チンプを使ってこの実験をすると、ある場合にはお互いにコインを入れあうようになります。どういう場合かというと、まずこちらのチンプにお金を渡します。次には、こちらのチンプにお金を渡します。順番にお金を渡してやるんですね。そうすると自分に渡されたお金を順番に入れていきます。これは一見お返しをしているみたいに見えるんですね。だけど、お互いにコインをいれあうようになるのは、人間が順番をちゃんとつくってやる場合だけなんですね。人間が交互にコインを渡してやると順番にお金を入れて、相手の入れたコインで自分が餌をもらう。餌をもらい合うということをします。

ところが、それぞれのチンプに一度にたくさんお金、コインを渡しちゃうとします。そうすると何が起こるか？それでもやっぱり人間の場合は1枚入れて相手が返してくれるまで待って、返してくれたらまた1枚入れようということをしますよね。だぶん。ところがチンプはそうしないんです。基本的にどっちか一方が入れ続けて、もう一方が食べ続けて、そのうちに入れ続けたほうが怒り始めるんですけども、そのフィルムを見ていると面白いですね。部屋の仕切りに空けてある穴から手を出してががーっと怒るんですけれども、それをやってもあんまり渡さない。

だからこれは、基本的に自分が渡さないと相手が渡してくれないだろうという理解ができないということなんですね。そういった理解をするように脳が作られていないんです。そういった形で理解できるようには心が働かない。それは単なる頭がいいとか悪いということだけとはたぶん違うと思うんですね。その話を始めると長くなっちゃうのではしょりますけれども。

同じ山本さんがやった実験でこういう

実験もあります。面白い実験です。ともかくチンパンジーは餌を前にするともう我を忘れてしまう。餌で頭がいっぱいになってほかのチンプのことを考えられなくなってしまう。だから食べ物じゃないものを使ったらどうなんだろうという実験なんですね。

手前のチンプは杖を持っています。奥のチンプの檻の前に餌が置いてあって、つえで引っ掛けると餌を取ることができるようになっています。

さて、手前のチンパンジーは、奥の部屋にいるチンパンジーにつえを渡してやるでしょうか？実は手前のチンパンジーの檻の前にはジュースが置いてあるんですが、ストローがないと飲めない。奥のチンパンジーにはストローを渡します。お互いにストローとつえと交換するでしょうか？こうしたことを調べる実験です。

この実験では、杖とストローを渡しあうということをわりとやるんですね、つえを渡してストローを渡してもらう。だけど自発的にそうするわけではない。何が起こるかというと、この写真にありますように一方が「よこせ、よこせ」と言うんですね。言葉ではなく身振りでですけど、よこせ、よこせと言うと渡すと。わざわざそんなことをしなくても相手がつえが必要だということは分かっているんです。別のやり方で確認してあります。だけど、相手が杖を必要としていることがわかっているだけでは、そのつえを渡してやらない。だけど要求されると渡す。このことは、実は私は、すごく社会科学にとって重要な発見だろうと思っていますね。

つまり、ここに社会規範の芽があるんじゃないかということです。自発的に渡すのではない。だけど要求されると渡す。なぜ要求されると渡すかというと、この実験では檻が仕切られていますから喧嘩にはならないんですけど、普通、要求されて渡さないと攻撃されてひどい目に遭うわけですね。だから攻撃されて自分が大変な目に遭うぐらいだったら渡しちゃった方がいいということで相手の要求に従う、そういう話なんだと思う。

だからこれは、お互いの長期的な利益を確保するための互恵性とたぶん違うだろうなという気はするんです。それは互恵性というよりは、社会規範のような、つまり規範に従わないと大変な目に遭っちゃうよということの一番の芽みたいなものだと思うんですね。それが１対１のときじゃなくてもう少し一般化してくると、社会規範になるんだろうなという気がしています。

こういった研究は、まだ話しだすといくらでも面白い研究はあるんですけれども、こういう研究を見ていくと、他の動物に比べて人間では互恵行動が非常に頻繁に見られるという特徴がある。よく見てみると、人間以外の動物でも互恵的に行動するように見えることはあるんですけれども、それはよく見るとそうだということです。先ほど紹介した霊長類研究所の松沢さんたちも、最近フィールドで、食べ物とセックスを交換するという話の例を見つけていますけれども、人間の社会でお互いに助け合う、あるいは互恵的に行動するということの頻度とはまったく比較になりません。

それではチンパンジーにはなぜ互恵的な行動がができないのか？１つは頭がいいかどうかという問題ですね。個体識別ができるか、あるいは記憶をちゃんと保持できるか、それから報酬遅延ができるか。そして、もっと重要なのは、ちゃんとした心の理論を持っていかどうかですね。相手の気持ちになって、相手がこうすれば相手はこう行動するだろうということが読めるかどうかです。

心の理論というのは、最近、心理学では非常に重要な概念として使われているんですけど、要するに他の個体が心を持っていて、その心に応じて行動が生まれてくるという理解、これは我々にとってはあまりにも当たり前のことなんですけれども、それはかなり高度な情報処理が必要です。

だいたい人間では、子供が心の理論を持ち始めるのは４歳から５歳にかけてだといわれています。今日はどうも時間がありそうなので余分な話もしちゃうんですけど、それはどういうことをやって調べるかというと、一番よく使われるのはサリー・アン課題という課題です。

サリーちゃんとアンちゃんという２人の女の子がいて、サリーちゃんが机の引き出しの中にチョコレートを入れて外へ出掛けます。その後にアンちゃんがやって来て、机の引き出しの中からチョコレートを取って、たんすの引き出しの方に移します。アンちゃんは外へまた行っちゃって、今度はサリーちゃんが戻ってくる。さあ、サリーちゃんはチョコレートをどこにあると思って探すでしょうという問題を子どもたちにやってもらいます。そんなもの我々にとっては、机の引き出しに決まっているじゃないかと思うわけですよね。

ところが３歳ぐらいまでの子供は、たんすの引き出しだと言うんです。つまり自分の知っていることと、ほかの人の知っていることが違うんだと

いう理解ができないんですね。サリーちゃんはチョコレートが机に引き出しの中にあると思うためには、ほかの人が知っていること自分の知っていることは違うと知っていて、そういった知っていることに基づいてほかの人は行動するんだという考え方ができる必要があるんです、そう考えるというのが心の理論があるというです。人間は５歳、６歳になるとほとんどできるようになります。多少の時間の差はありますけれども。

人間的な意味での互恵行動をするためには、この心の理論が絶対に必要なんですよね。今、助けてやると相手は自分が困ったときに助けてくれるだろうと思えないといけない。もちろんそのことが分かっただけでは、別に互恵行動を取るわけじゃないですけど。

それが分かるのは一種頭がいいからですが、だけど頭がいいというだけでは不十分です。世の中にはゲーム理論家というとても頭がいい人たちがいます。そういう人たちが互恵的に行動するかというと、そうではない。彼らにとっては、ばかなやつは互恵的に行動するんだというわけですね。だからそれは頭がいいだけでは互恵行動は生まれないという一番いいその証拠だと思っています。

互恵行動が生まれるためには、もう１つ条件があって、非血縁関係にある他個体と、長期的な関係を結んでいるということが、たぶん非常に重要なことですね。そういう環境の下では、互恵的に行動した方が結局は長い目で見て自分の利益になる。そういう環境があって、しかも心の理論を持っていることが大切なんだろうと考えています。だから、こうしたことは人間の社会性を考えるときにとても重要なことであると思います。

利他行動の基盤

こうした互恵行動は人間の利他的な行動を考えるときに、たぶん非常に重要な意味を持つことになります。いわゆる利他的な行動の多くは、実は互恵行動であると考えます。通常の心理学では、自分に利益が絶対戻ってこないときにコストを掛けて他者に利益を与える行動を利他行動と言っているんですけれども、そういう行動に見える場合も、もっと長期的な観点から戦略として考えると、それはたぶん互恵行動なんだろうと考えています。

実際、ヒト以外の動物では血縁者への利他行動を除くと利他行動はほとんど見られません。もちろん社会性昆虫は話が別です。なぜ社会性昆虫が利他行動を取るかというと、それぞれの個体が自分の遺伝子をそのまま自分の子孫に残せないからです。だから利他行動を取るような個体がいたときに、そのことによって自分と同じ遺伝子が残る確率が減るわけじゃないですね。結局その遺伝子が残るためには女王がいて、その女王と共有している遺伝子を残すしか道がないので、当然そういう利他行動は起こります。

ところが人間をはじめとする哺乳類は、それぞれの個体が直接その子孫を残すわけですから、自分の損になって、ほかの人の利益になるような行動を取れば、それは相対的に自分の利益が減って、自分と同じ遺伝子を共有する子孫が減ることになるはずです。だからそういう特性は進化し得ないはずなんですよね。

しばらく前までは、しばらくといってもずいぶん前なんですけど、我々が中学生、高校生だったころには、個体の淘汰と種の淘汰という概念が、確か中学校、高校で教えられていたと思うんですけど、そんなものはもうごみ箱行きになっています。種の淘汰というのはあり得ない。淘汰が起こるのはすべて遺伝子のレベルで起こる。

ただし、種の淘汰じゃなくて集団淘汰となると、これは少し話が違います。集団淘汰説というのは完全に否定されたかというと、必ずしもそうではない。ある特定の遺伝子を共有している集団が、例えばお互いに協力するという遺伝子を共有している集団はうまくいくから、そういう性質を持っている集団が拡散していって、そういった性質が進化するのではないかという説が集団淘汰説というんですけれども、これは普通の生物学の中では理論的には可能ですが、ほとんど起こらないと考えられています。

集団間の淘汰圧と集団内での淘汰圧の力関係によって、場合によっては集団淘汰は起こることがあり得るんですけれども、それは非常にまれであると考えられています。なぜかというと、通常の場合は集団内での淘汰圧の方が集団間の淘汰圧より大きいからです。つまりいくらみんなうまくいくような性質を持っていても、うまくいく集団の中で自分だけただ乗りする人がいると、そちらの方がもっとうまくいく。そうすると、そのうまくいっている集団の中で、ただ乗りする人間が増えていきます。

ですから集団の中では自分勝手なやつがどんどん増えていって、集団間

では集団のために協力する性質が増えていくんですけども、だいたいの場合は集団内での淘汰圧、つまり集団内で自分勝手なことをする遺伝子の拡散の方がずっと早い。だけど、ある種の人たちは、特に一部の経済学者と人類学者が信じているんですけど、人間の場合には集団淘汰はありなのだということを言っています。

それはなぜかというと、人間は制度をつくって集団内での淘汰圧が低くなるような仕組みを作っちゃったからだと言うんですね。例えば一夫一婦制がそうです。一夫一婦制があると、その集団の中で子孫を残す分散が小さくなる。一夫一婦制の下で100人も子供をつくることはできないですね。そうすると集団の中でうまく立ち回った人間が、一生懸命まじめに働いた人間よりもたくさんの子孫を残すことになったとしても、それほど極端な差が生まれることがなくなります。そのために集団間での淘汰圧が、より相対的に大きく働くようになる。そういうことを言う人もいます。

そういう話なんですけれども、それでは、人間以外の例えばチンパンジーで利他行動というのはあるんだろうか、ということになる。これも面白い実験がいっぱいあります。一番有名なのはジョアン・シルクという人がやった実験で、こういう実験なんですよね。チンパンジーが２人います。こちらのチンパンジーは２つのオプションを持っています。こっち側のオプションを引っ張ると食べ物がそれぞれのチンパンジーに行きます。こちらのオプションを引っ張ると、自分だけ食べ物がもらえます。さあ、どうするでしょう。

もし相手のチンパンジーのことを考える心があれば、当然二人とも食べ物をもらえるオプションを選びますよね。あるいは将来、付き合いのことを考えると、「お前、あんな勝手なことをやりやがって」と文句を言われるといけないので、やはり二人とも食べ物をもらえる側を選ぶ。逆に、意地悪な人はこっち側を選ぶはずなんです。「お前、悔しいだろう」ということを見せつけることが愉しいというひとは、自分だけ食べ物がもらえる方を選ぶはずです。だから、どちらかに偏るはずなんですよね。あるチンプは二人とも食べ物をもらえるオプションを選ぶけど、別のチンプは自分だけ食べ物をもらうオプションを選ぶということはあるかもしれない。だけど、この実験をやってみると完全にランダムなんです。もう一人のチンプが食べ物をもらえてももらえなくても、「そんなことは知るかよ」と全

然気にしない。

この実験からは、チンパンジーにはある意味では利他性、相手のことを考えてする行動というのは、やっぱりそういう心の働き方というのはないんだろうとも考えられます。これに近い実験は我々も子どもを使ってやっているんですが、何歳ぐらいからそうなるのかというと、これもだいたい心の理論が出来上がってくるころですね、概していうと。

ちょっと前置きが長くなってきたのではしょりますけれども、どうして人間は利他性を持っているのか？さっきの食べ物の実験の例でも、人間だったら、だいたい自分が損をするわけじゃないんだから、相手にも食べ物をやった方がいいじゃないかと思うと思うんですけども、よっぽど意地悪な人じゃない限りですね。だからそういう意味で、たぶんチンパンジーと人間では頭のよさとは全然別な意味で、基本的な心の働き方の違いがあるんだろうと考えられます。

人間はどうしてそんな利他性を持っているんだろうかということに関しては、いろいろな説明があるんですね。よく言われるのは互恵的利他性、これは互恵性の話ですよね。人間以外には互恵性だってあんまりない。それから血縁淘汰の話があります。要するにそれは血縁でつながっている、遺伝子を共有している他個体を助けるのは自分自身を助けるのと同じことだということで、これでも説明できるんですけれども、例えばいま紹介した食べ物の実験を人間でやれば、相手が自分と血縁でつながっていなくても利他的な行動は取ると思うんです。

どうして人間は違うんだろうというときに、１つよく使われるのが、規範の話です。人間は規範を内面化しているんだという話し。これは、現象の記述としては、たぶん正しいですよね、そうしないといけないと思っているんだから。だけど、なぜ人間はそうしないといけないと思うようになったんだろうと考えると、それでは説明できなくなってしまうんです。

これは非常に議論のあるところだと思いますけれども、血縁淘汰と互恵的利他性だけでは人間の利他性の説明としては不完全だということになると、それ以外の説明を考えないといけない。実際、いくつかの説が提案されています。

簡単に話すと、１つはtolerated theft、つまり強奪されるのに抵抗しない

で、先回りして攻撃してきそうな相手に資源を渡してしまうという考え方があります。食べ物などの資源を自分だけで独り占めをしようとすると、まわりの他個体からの攻撃を受けて結局は損をするからだということです。

たとえばある猿がたくさんの餌を持っているときに、他の猿が俺にもくれと寄ってくる。嫌だねと突っぱねる。相手が一匹であれば突っぱねることもできるけど、何匹もの猿が俺にもよこせと来る。すべての猿と争っていると、結局大変な目にあってしまう。こんなたくさんの連中と争うのはばからしいから、いちいち争わないで、「自由に取っていってください」と言うというのが、tolerated theft による利他行動という考え方です。だから、ほんとうは利他行動じゃないんですが、行動だけ見るとまわりの他個体に利益を分け与える利他行動のように見える。

実際にさっき紹介した山本さんの実験で、チンパンジーが要求されると要求に従って杖を渡したというのは、これに近いんだろうと思うんです。ある程度、先読みですね。自分で独り占めをすると大変な目に遭ってしまうというのを先読みするから、持っていっていいよと言うと。

それから重要な要因が共感性ですね。つい共感してしまう、相手のことがかわいそうになってしまう。これは非常に重要なんですけれども、なかなかいろいろな説があって考えるのが難しい。これに関しては、また全然違う系統の実験があるんですね。

これはもうずいぶん古い、1964年の実験なんですけれども、2匹のアカゲザルを別のおりに入れておく。一方の猿には電気ショックがいくようにしてあります。もう一方の猿には鎖を引くと餌が来るようになっています。だけど鎖を引くともう一方の猿に電気ショックがいきます。ものすごく強烈な電気ショックがいきます。だからもう苦しんで泣き叫ぶわけですね。そのときに本当にほかの猿のことなんか知らんよというんだったら、鎖を引いて自分で餌を取っちゃうはずなんですね。

ところが、よっぽどおなかが減ってくるまでは、なかなか鎖を引かないんです。１週間ぐらい引かなかった。本当に自分が飢え死にするようになるまでは鎖を引かない。さっきのバナナの話と全然違うんですよね。たぶんこれは全然違うモジュールが働いているんだと考えられます。だから積極的に助けるということと、自分の仲間を苦しめることは避けるということと、私たちは同じ利他性という言葉で言うんですけれども、たぶん全然違う性質のものだと考えられます。

どうして人間は利他性を持つかという理由としては、こう言った意味での共感性が考えられますね。なぜ利他行動を生み出すような、言い換えれば自分が損をしてまで他個体に利益を与える行動を生み出す心のはたらきである共感性がヒトにはそなわっているかといえば、そうした共感性が有利に働く場面があるからだと考えられます。もちろん子供との関係で共感性が生まれてくるということは十分に考えられますが、それはまた別の話です。子どもとの関係以外の場面でなぜ普通の人が他人に対して共感するかというと、たぶんそれは、共感性を持つことで、ほかの個体の行動が読めるようになるからだと考えられます。つまり自分がその場になったらどう思うだろうかという形で、自分の中にそのときに起こる感情をシミュレートする。そうすると相手がどう思うだろう、どう行動するだろうということが分かるようになります。

共感性を通して他者の感情をシミュレートするやり方の特徴は、認知的なコストをかけないで他人の行動を予測できるという点にあります。他人の行動を予測するもう一つのやり方は、理論的に考えて、こういうときにはこういうタイプの人間は相手はこう行動すると、こう思うだろうと考えるやり方で、そのためには認知的なコストがかかります。共感性を通して他者の行動を予測するときには、そうした認知的なコストをかけなくても周りの人の行動を予測することができるようになるんですよね。だけど、そのためには、それなりの別の種類のコストが掛かります。

それはどういうコストかというと、自分の中に感情をシミュレートしてしまうので、そのことによって自分自身の行動が制約されちゃうということですね。ヒトは常に相手の感じることを自分の中にシミュレートしています。こうした感情のシミュレーションについて、最近ミラーニューロンという脳のニューロンが、非常に重要な働きをしていることがわかってきています。例えば、自分がある行動を取るときに働くのと同じニューロンが、ほかの人が同じことをしているときにも働いてしまうということがわかってきたんですね。自動的に働くようになっているんですね。

だからそういう形で自動的にほかの人の気持ちを、その細かいメカニズムというのは、まだ全然分かっていないんですけれども、そういう自動的に感情をシミュレートするようなことをすると、非常に効率よく周りの人

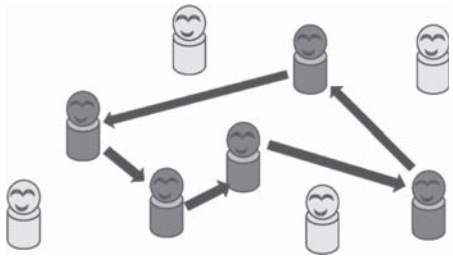
間の行動を予期することができます。そのために他人の行動を予測できて適応的な行動を取れるんですけれども、だけど、そうすると、ひどいことができなくなってしまうというコストがかかることになります。

ひどいことをすると、相手が苦しむ感情が自分の中に起こってしまうので、そういうことをやりにくくなると。だから他人を苦しませることがわかっている行動をとれなくなってしまう。これが、サイコパスと呼ばれているような人たちだと、そういうことが起こりにくいんですね。そういう人たちは、たぶん他人の行動を読むときに感情のシミュレーションとは違うストラテジーを取っているんですね。感情をシミュレートするんじゃなくて理論的に詰めていくんです。

だから、例えば結婚詐欺をするときに、こういうことを言えば喜ぶということを理論的に分かっているんですね。自分がそう言われたらうれしいだろうと思ってするんじゃなくて、この手の人はこう言えば喜ぶといった理論に従って、相手が喜ぶようなことを言う。だから自分の中に感情が起こらない。だから自分の中に起こった感情にとらわれて、つい手を抜いてしまうということをしないで自由にひどいことができる。

それほどじゃないんですけれども、経済学を履修してゲーム理論についてしっかり勉強すると、心理学を履修した学生に比べて利他的な行動を取らなくなるという実験もあります。北大でやったら分かりません。学生がそれほどまじめに勉強していないから。こうした実験を経済学部の学生を使ってやってみると、その大学で経済学の学生がどれぐらいまじめに勉強しているかが分かります。もちろん冗談ですけど。

ヒトの利他性の説明の３つ目が、間接互恵性です。申し訳ありません、これまでにずいぶん時間を使っちゃったんですけれども、私が今一番中心的に研究しているのがこの間接互恵性という話です。間接互恵性というのは要するにどういうことかというと、みんなが他人を一方的に助けている状態です。ただし、他の人を助けている人しか助けない。そうすると、この図にあるように、他の人を助け



ていない人は別の人から助けてもらえない状態になってしまいます。こうした状態で、他人を助けるためのコストよりも、別の人から助けてもらうことで得られる利益の方が大きい場合には、みんなすすんで他人を助けるようになるはずです。

つまり、援助は一方的ですから、助けた人から直接お返しが来るわけじゃないんですけれども、回り回ってお返しが来る。「情けは人のためならず、まわりまわって己が為」というわけですね。これが間接互恵性といわれているものです。利他主義者であるという評判を取っていると、ほかの人たちから助けてもらえるようになる。人々がそういう行動形態を持っているときには、他人を助ける人の方が、利己的に行動する人よりも、より多くの利益を得ることができる。ただし、この図は非常に単純化されていますけど、そういうことが起こるためにはいろいろな条件が必要です。

こうしたかたちで間接互恵のシステムが成立する条件を特定する作業は、数理モデルを使ったり、コンピューターシミュレーションをしたりして、今いろいろな人たちが研究しています。例えば数理生物学者が研究しているし、数理人類学者もやっていますし、経済学者もそういう研究をしています。心理学者の中にも、そうした理論モデルの研究をしている人がいて、我々のCOEの中では高橋伸幸さんが、この理論モデルの研究をしています。

間接互恵性の心理学的含意

私自身は、この間接互恵性に非常に興味を持っているんですけれども、私の関心の持ち方は、どういう理論的なモデルが成り立つかということじゃなくて、もう少し心理学的な関心です。こういうことが基になって人間の利他性が進化してきたのであれば、人間は評判への非常に強いセンシティビティーを持っているだろうと予想されるので、実際にそういうセンシティビティーを持っているかどうかを確かめています。そういう研究をしています。それが今日お話ししようと思っています研究です。

ということで、ここまでは前置きで、これから本番なんです。時間を取っちゃって申し訳ありません。実はここからは、ほかのところで話した話をそのまま使わせていただくので、日本語じゃなくて英語のスライドを使

わせていただきます。申し訳ありません。

これは実は去年、オーストリアのウィーンであったわりと小さなワークショップで話したんですが、そのワークショップは半分が数理生物学者で、ほぼ半分が経済学者で、ノーベル経済学賞をもらった方が２人来ていたという、結構インテンシブなワークショップだったんですけれども、小さなワークショップですね、そこで話した話です。多くはその数理モデルの話をされたんですけれども、たぶん評判に対するセンシティビティーについての研究をやっている人は、そこでは私だけですね。

間接互恵性の話は、日本語で言う「情けは人のためならず」という話です。もし実際に、情けは人のためならずという状況が世の中に存在しているとしたら、それが成り立つためには、人間は評判に対するセンシティビティーを持っていないということになります。

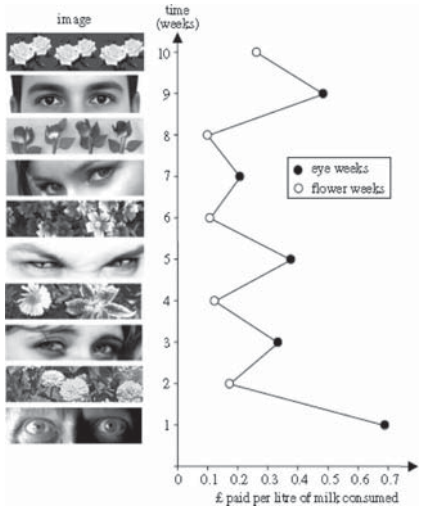
その評判に対するセンシティビティーには２種類のものがあります。まず、ほかの人についての評判を気にするという意味でのセンシティビティーがあります。ほかの人が、いいやつなのか、悪いやつなのかを見分けるために、他人の評判を気にしている。この点に関しては、進化心理学者のジョン・トゥービーとレダ・コスミデスという人たちがずっと研究をしています。

私の関心は、もう１つの方ですね。もう１つのセンシティビティーというのは、自分が周りから見られているということに対するセンシティビティー。ほかの人が、いい人か、悪い人かということに対してセンシティブになると同時に、自分が変なことをやると、みんな見ているんだよ、だからそういうことに気を付けないといけないねという心の働きを、ほぼ自動的に持っているんだろうと考えられます。要するに見られていると、みんな協力的に行動する。

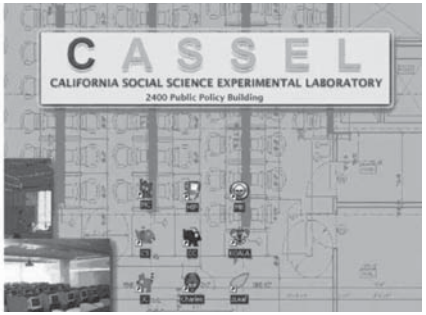
実は本当に見られていなくても、自分がモニターされているということを示す手掛かりがあれば、それだけで人間はつい利他的に行動するようになるということが知られています。実際に見られていなくても、です。つまり、それはどういうことかといいますと、人間の心はかなり自動的に働きますから、ある手掛かりが与えられると、それに応じたモジュールが活性化する。そのために、そういう状況で適切な行動が取られるようになる。

１つ面白い、これは私が大好きな実験なんですけれども、ベイトソン、ネットル、ロバーツの３人がやった実験です。これはどういう実験かというと、皆が共同で使っている研究室のラウンジを使った実験です。研究室のラウンジの中には、だいたいどこの研究室でも、コーヒーマシンが置いてあるわけですよね。みんなでお金を出し合ってコーヒー豆を買ってきて、それを使ってコーヒーを入れて好きなように飲むことができる。お金を入れてくださいという箱だか缶とかがあって、そこに自発的にお金を入れて自由にコーヒーを飲みます。お金を入れているかどうかを誰かが見はっているわけではありません。

そのときにコーヒーマシンの背後の壁に、花のポスターと目のポスターを、１週間置きに張り替えておくんです。そして、毎週コーヒーを飲んだ人たちがいくら払ったかを調べます。コーヒー１杯当たりいくら払ったかを調べるんですね。こういう研究は面白いですよね。コーヒーマシンの背後にこの図に示されているように、目のポスターがあると、お金をちゃんと払うようになるという結果でした。コーヒーを飲んでいる人たちは、実際に見張られているわけでも何でもないんですね。だけど、こういう目のポスターがあるだけで、それが手掛かりになって、間接互恵性を生み出すための特定のモジュールが活性化する。すると、それに応じた行動が出てくるということです。



もう一つの実験を紹介します。ヘイリーとフェスラーという人がUCLAでやった実験です。これは基本的にはディクテーターゲームの実験です。ディクテーターゲーム、日本では独裁者ゲームというんですけども、これは非常に単純なゲームで、被験者を２人連れてきて一方にお金を渡します。そして、自由に２人の間でお金を分けてくださいと言う。実験の参加者に



実際に分けてもらいます。その結果、その人が分けたお金を二人の参加者がそれぞれ実験参加の報酬として手に入れて帰ります。

そのとき実験の参加者は、お互いに相手が誰なのか分かりません。実験者にも、誰がどのように分けたか分からないように設定しています。だから、実験に参加した人は、自分がどう行動したと思われるかをまったく気にしなくても済むようになっています。

そのときに、実験に参加した人の目の前にコンピューターを置いておいて、そこにこの図の左側に示されているコンピュータ画面の壁紙か、これはエジプトのホルスの神の絵なんですけれども、右側に示されている壁紙か、どっちかを表示しておく。そうすると相手に渡す金額が10ドルのうち、目の方があるときには3ドル14セント、目がないときには2ドル83セントと、目の前のコンピュータの壁紙に目が表示されているだけで違うんですね。

次に紹介するのは、リグドンという人たちがやった実験を基にして、我々自身がやった実験です。この実験は、囚人のジレンマを２人の参加者の間でやってもらう実験です。

この実験では、コンピュータの画面ではなく紙に印刷した囚人のジレンマについてのインストラクションを読んでもらって、協力の手を選ぶか非協力の手を選ぶか（実際の実験では「協力」とか「非協力」という言葉は使っていませんが）を決めてもらいます。そのときに紙の上にこういう点

を３つ打っていく。これは実は顔に見えるんですね。あるいは、この３つの点を上下反対にした点を打っておく。これは顔に見えない。だけど、この実験に参加している人たちは、

こうした点を見て、「あ、これは顔に見えるぞ」とか、「こりゃあ顔には見えないや」とか、そんなことを考えてもいないでしょう。そんなことを意識的に思っているわけじゃない。だけど、こうした逆三角形の点が３っつあるか、正三角形の点が３っつあるかで、囚人のジレンマでの協力度が違うんです。逆三角形の点が打ってあると協力するようになる。だからまったく意識には上っていないんですけれども、こういうものが手掛かりになって、ここは協力する必要があるんだということを自動的に決めちゃっているんですね。

しかもこれが面白いのは、囚人のジレンマの実験の後で、3つの点を印刷した紙を渡して、これらの点を使って絵を描いてくださいと言うんですね。何でもいいから絵を描いてください、と言います。そのときに顔を描いた人、それから家とか何か道具、物を描いた人、それからこれは単なるデザインを描く人とか、いろんな人がいます。こうやって絵を描いてもらうと、３つの点を使って顔を描く人は、囚人のジレンマでよく協力する傾向があります。だからたぶんそういう人は、自分が見張られているということに対して、モニターされているということに対して非常にセンシティブな人だと思うんですね。

逆三角形の三つの点を見ると協力してしまう人は、ほんとうはただ乗り可能な状態で協力してしまうわけなので、遺失利益を生んでしまっていますね。だから、そうした人は自分の利益をちゃんと確保できない愚かな人ということになりそうですが、ほんとうにそうなんでしょうか？だけど私は、そうした人がほんとうにバカな人なのかどうかは一概に言えないと思っています。というのは、評判にセンシティブな人はただ乗り可能な場面、つまりほんとうは見はられていない場面でも協力してしまい、非協力行動（ただ乗り）から得られるはずの利益を失ってしまう。これは自己利益の最大化をはかるという点から見ると、エラーを犯していることになります。つまり、馬鹿なことをしているわけです。

しかしこうした場面では、別の種類のエラーを犯してしまう可能性について考えておく必要があるのではないでしょうか。つまり、ほんとうは見はられているのにずるい行動を取ってしまい、「あいつは自分勝手な奴だ」という評判を立ててしまうというエラーです。

こうした二種類のエラーというのは、統計学で言う第１種のエラーと第

２種のエラーのようなもので、どちらがより重要なエラーなのかを一概に言うことはできません。例えば統計的な推測で第１種のエラーと第２種のエラーのどちらを小さくすべきかは、調べている現象に依存します。ある薬がある病気に効くかどうかを調べているときには、ほんとうは効かないんだけど効くという結論を下してしまうのが第１種のエラー、ほんとうは効くんだけど効くという結論を下さないのが第２種のエラーです。効果のない薬を認可するのは問題だという政府の立場からは、第１種のエラーをなるべく小さくすることが重要です。しかしこれが、副作用があるかどうかを調べているのだとすれば、ほんとうは副作用があるんだけど副作用がないという結論を下してしまうと、そのために多くの人が副作用に苦しむことになります。だから、なるべく第２種のエラーをへらさないといけない。そのために第１種のエラー、ほんとうは副作用がないんだけど副作用があると結論してしまうエラーは多少がまんしないといけない。

それと同じことが、評判に対するセンシティビティーにも言えるだろうと思います。評判に対するセンシティビティーが高い人は、見張られていないときに見張られていると思う方にエラーをたくさん起こす。それで損をすることもあるんですけれども、逆方向のエラーを起こす可能性が小さくなる。つまり、実際に見張られているのに見張られていないと思ってしまうエラーを起こしにくくなる。このどちらのエラーの方が重要な結果を生むかによって、どちらのエラーを小さくすべきかが決まってくるはずです。そして、どちらのエラーがより重大な結果を産むかは社会によって違うと思うんです。

多くの場合は、悪いことをやっているのがばれちゃったときの方が、本当は悪いことができるんだけれども、それをしなかったときのコストより大きいですね。悪いことがばれると集団から追い出されてしまったりして、生きていけなくなってしまう。集団から追い出されたときに、もう行き先がないような状況だと、それはすごく困ったことです。だからそうした状況では、どちらかのエラーを起こすとすれば、見張られているというふう思い込むにエラーを起こしちゃった方がいいだろうという意思決定の仕方をするようになると考えられます。たぶん人間は、そういう形である種のバイアスを組み込んで状況を判断していると考えています。

評判がヒトにとってとても重要な役割を果たしてきたというのは、人間

というのは、だいたいそもそもがゴシップ好きで、ゴシップにものすごくセンシティブだということによくあらわれています。いろいろな会話の分析なんかをしてみると、だいたいもう人間の会話のほとんどがゴシップなんです。あの人はどうしている？あの人のことは知っている。あの人は最近こんなことをやっていたなんて、面白いね。ええっ、本当にと、こう、すぐ乗ってきますよね。だから人間は、そういう形で非常に頻繁に評判情報を交換しているんです。

特に自分の集団の中でそういうことをする。ほかの全然関係ない人の評判なんか知ったってしょうがないわけですね。だからそういう全然知らない人の評判というのは、そんなことを言われたって、へーっと言うだけでまったく面白くない。だけど自分の知っている人が、ええっ、こんなことをやったのというと、もうすぐみんな集まってきて聞きたくなる。だから自分の集団の人のやることに対して非常にセンシティブです。だからほかのそういう自分と関係ないところに行くと、旅の恥はかき捨てということになって、恥ずかしいことを平気でできる、ということになります。

最小集団実験と、社会的アイデンティティー理論 vs. 評判理論

こういう、本当に人間は自分の集団の中で評判を気にするんだろうかということを調べる実験をいくつかやったんですけれども、これからお話しする実験の元になっているのが、最小集団実験という実験です。これは1970年ごろの初めから社会心理学者が実施してきた実験で、一時、今でもそうなのかもしれないんですけれども、集団間のコンフリクトの説明として重要な理論的役割を果たしてきた社会的アイデンティティー理論の基になっている実験ですね。

社会的アイデンティティー理論というのはどういう理論かといいますと、要するに、人間は自分のアイデンティティーの一部を集団に依存しているんだ、例えば自分は日本人だということは、自分自身のアイデンティティーの中に組み込まれて、集団へのアイデンティティーが、集団のカテゴリーが自分のアイデンティティーに組み込まれてしまうという基本的なアイディアに基づく理論です。

だから、自分の集団に対しても、自分自身に対するのと同じような形で

行動することになります。さらに、社会的アイデンティティー理論の信奉者たちは欧米の社会心理学者が中心になっていますから、人間は自尊心を高く保とうとする傾向を持っているという基本的な前提を受け入れています。だから、自尊心つまり自分は他人よりも優れているという信念を保とうとする傾向を持つ人間は、自分だけではなく自分が属する集団も他の集団より優れていると考える傾向があるということになります。これが社会的アイデンティティー理論です。極端に言えば、人間はほかの集団との差をつけることで自分自身が偉くなったような気になる。要するに自尊心を上げるために集団の間に差をつけたがるんだという理論です。最近は理論の内容がちょっと変わっていますけれども、基本的なアイディアはそういったものです。

私に言わせると、おばかさんな理論ですね。この理論によれば、別に集団間で本当に利益の対立がなくたって、集団間にコンフリクトが生まれるんだと言っているわけですよね。それはどうしてかというと、人間は集団テゴリーとの間に自分のアイデンティティーを求めるので、つい自分の集団がいいと思う。自分の集団を優遇したくなる。だから自分の集団と比べて相手の別の集団を押し下げたくなると、そういう心を持っているからと言うんですよね。

なぜそんな理論が出てきたかというと、その一番基になっているのが、この最小条件集団実験なんです。これはどういう実験かといいますと、まず実験参加者をささいな基準によって２つの集団に分けます。よく使うのがクレーの絵とカンディンスキーの絵とどちらが好きかで集団を分けるやり方です。何枚もクレーの絵とカンディンスキーの絵を見せて、どちらが好きですかと聞いていって、はい、あなたはクレー好きの集団の人です。あなたはカンディンスキーファンの集団の人ですねと、集団を２つに分ける。多くの場合は実際にどちらの画家の絵が好きかで分けるのではなく、ランダムに分けておいて、あなたはクレー集団、あなたはカンディンスキー集団と告げます。クレーの絵とカンディンスキーの絵を区別できる参加者はほとんどいないと思ってるんですね。ちなみに私はクレーの方が好きです。

そうして参加者を二つの集団に分けておいて、一定の報酬を自分の集団の人１人と、もう１つの集団の１つに分けてくださいと参加者にお願いする。参加者には、二人に分ける金額に差をつける理由がまったくないのに、つい自分の集団の人にたくさん渡して、相手の集団に少ししか渡さないという行動を取ってしまうというのが、この最小条件集団で多くみられる結果です。だから、実際に何の利益の葛藤もなくても、集団間に差をつける心の仕組みを人間は持っているんだという、社会的アイデンティティー理論の説明がもっともに思えてしまいます。

こうした実験を基に社会的アイデンティティー理論というのをつくって、集団間の対立というのは、人間の本性に組み込まれているんだという説明をする。社会心理学者はそこまでは言わないんですが、進化心理学者の中には最近はそういうことを言う人もたくさん出てきました。これはちょっと困ったことだと思っています。

これもずいぶん昔の話になるんですけど、90年代の初めころですが、こうした最小条件実験の結果の解釈は違うだろうと考えて、それは違うよということを証明するための実験をやりました。社会的アイデンティティー理論が受け入れられたのは、カテゴリーだけの実態のない集団を使っても集団間の差別行動が生まれるという実験結果が得られたからですが、私たちが証明したのは、最小条件集団には何の利益の葛藤がないと思っているけど、実はそうした利益の葛藤が存在しているんだよということです。最小条件集団実験においても集団間の差別行動が生まれるのは、実は、集団とのアイデンティティーではなくて、評判、特に自分の集団の人たちに対する評判についてのセンシティビティーが生み出している現象なんだということを示すための実験をおこないました。

アイデンティティー理論の立場から言うと、自集団を優遇し、相手の集団を貶めるのは、要するに、自分が優れていると思いたいという人間の基本的な性質のためだということになる。自分の集団は自分の一部なのだから、自集団を優遇することは自分を優遇するのと同じことになる。相手の集団を貶めれば、相対的に自分の集団が優れていることになる。そうすると、例えば囚人のジレンマゲームでは、囚人のジレンマの相手が自分と同じ集団の人の場合と、違う集団の人の場合で、協力する程度に差がつくという予測が生まれます。この集団というのは、さっき紹介したクレー集団、カンディンスキー集団といった、まったくどうでもいいような集団の場合でもそうなるという予測です。

ところが実は、そういうアイデンティティーの話じゃなくて評判の話なんだよという枠組みから言うと、自分が同じ集団の相手と囚人のジレンマをプレーしているか、違う相手と囚人のジレンマをプレーしているか、それ自身は重要じゃないということになります。問題なのは、相手が自分のことを同じ集団の相手と思っているかどうかである。そういうことを示す実験をしました。

つまり、相手が自分の集団の人間であるから自動的に自分の集団を優遇するのではなくて、相手が自分に対して、自分の集団の人間を優遇することを期待しているという理解があるということです。そうした理解があると、自分の集団の仲間を優遇しないと、あいつはけしからんやつだと思われてしまうのじゃないかと心配になるわけですね。だから相手が自分のことを自分の集団だと分かっているときには、優遇しないと仲間としてけしからんとい言われてしまう可能性があるので優遇するけど、そうじゃないときには、つまり相手が自分の集団の仲間でも、相手がそうは思っていない場合には優遇しないのであると予測されるわけです。この予測が成り立つかどうかを調べるために、２つの条件をつくりました。共有知識条件とプライベート知識条件です。

共有知識条件というのは、お互いに自分も相手も、それぞれがどちらの集団の人間かを知っている。相手も自分の相手は同じ集団の人間だということを知っている。しかも相手がそのことを知っているということを、プレイヤー本人が知っている。これがコモンナレッジ条件、つまり共有知識条件です。これまで行われてきたほとんどの最小条件集団実験は、こういう状況で行われています。

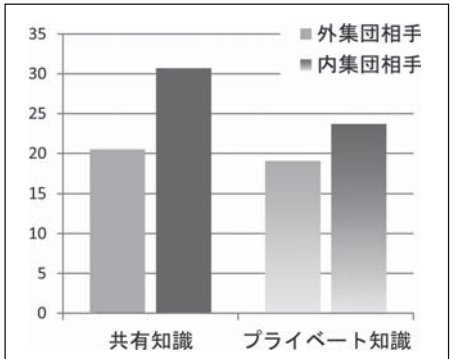
それに対して、プライベート知識条件というのをつくってみます。プレーヤー本人は、相手が自分の集団の人間だということを知っています。だけど相手は、プレイヤーがどちらの集団の人間なのかを知りません。そしてプレイヤーは、そのことを分かっています。つまり、プレイヤーは、相手が自分のことをどちらの集団の人間か知らないということを分かっています。これがプライベート知識条件です。

このプライベート知識条件だと、要するに自分が相手に何をしたって集団の中での評判にはつながってこないわけですね。だから旅の恥はかき捨て状態です。そうすると、内集団に対する協力度が上がるのがレピュテーション（評判）に対する心配があるからだとすれば、プライベート知識条件ではそれはなくなってしまうはずです。そうすると、この相手がどちらの集団の人間であっても差が生まれないだろうという予想になります。

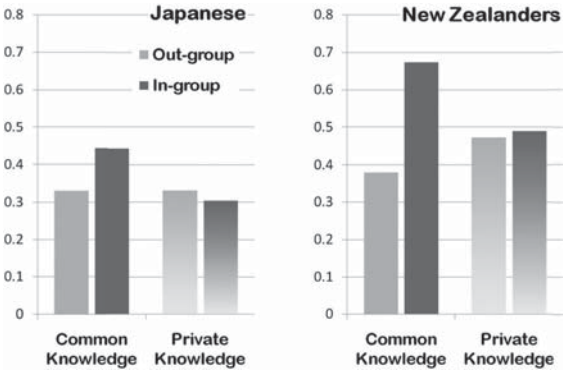
それに対してアイデンティティー理論の考えだと、プライベート知識条件でも、集団の間に差があるはずだという予測になります。相手が自分のことを集団と知っているかどうかは、自分が属している集団とのアイデンティティーとは関係ない。自分と相手が同じ集団かどうかというだけしか関係ないはずです。実際にこれは最初にやった実験なんですけど、コモンナレッジのときには相手がイングループかアウトグループかで非常に大きな差があります。

この図は、1997年に発表した論文での結果です。それ以前にも別の形の実験をやったんですけども、そちらの実験については省略させてもらいます。予測通り、共有知識条件では、相手が内集団の人間の場合と外集団の人間の場合で、相手に対する提供額に大きな差が見られています。これに対してプライベート知識条件では、内集団を相手にした場合と外集団を相手にした場合でちょっと差があるように見えますが、この差は有意な差ではありません。

ただし、プライベート知識条件でも本当は差があるんじゃないのと言われるといけないので、同じような実験を繰り返しやってみました。そうすると、実験のやり方がうまくなってきて、非常にきれいに結果が出てきます。つまり、プライベート知識条件では、内集団相手の場合と外集団相手の場合で、相手への提供額にまったく差が見られなくなります。



囚人のジレンマゲームで、100円の元手のうち相手に提供した金額。神信人・山岸俊男（1997）. 社会心理学研究 12, 190-198.



囚人のジレンマゲームでの提供率（元手の中から相手に提供した金額）。Yamagishi, T., Mifune, N., Liu, J. H., & Pauling, J. (2008). Exchanges of group-based favors: Ingroup bias in the prisoner's dilemma game with minimal groups in Japan and New Zealand. *Asian Journal of Social Psychology*, 11(3), 196-207

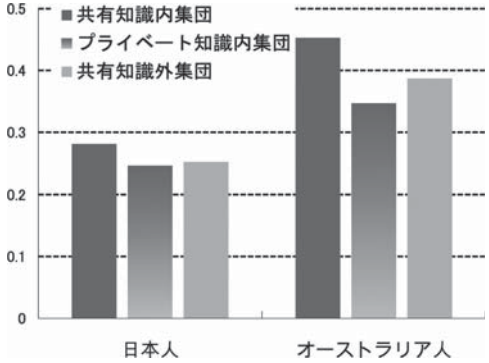
また、同じ実験をニュージーランドと日本でやりました。そのときも共有知識条件では、どちらの国も内集団と外集団の間に差をつけるんですけれども、プライベート知識条件だとまったく差をつけません。それは日本人もニュージーランド人も同じなんですね。ちなみにこういった実験をすると、日本人はほかの国の人たちと比べて協力しません。これはいろいろな形でやってもだいたい一貫しています。

一般に日本人は協力的だと思われているんですけど、こうした実験結果を見るとそれは違うんですよね。だいたいアメリカ人とかオーストラリア人とかニュージーランドとか欧米の人たちの方が協力率が高いです。日本社会で人々が良く協力しあっているのは、相互に監視し合う日本的なシステムが確立している状況の中で、そうせざるを得ないような状況をつくられているからなんですね。

だけど、こういう実験室のように、そういった日常生活の中に組み込まれている仕組みを取り去った社会的な真空状態をつくると、日本人は協力しなくなるということだと思います。いずれにせよ協力率の全体の高さ、低さは日本とニュージーランドでは違うんですけれども、相手が内集団か外集団かで差が出るのは共有知識条件でだけだということに関しては、日

本人もニュージーランド人も同じです。

次の図は、最小条件集団じゃなくて国籍を集団として使って日本とオーストラリアを比較した実験の結果です。日本人が参加者の被験者の場合には相手が日本人の場合が内集団、オーストラリア人が相手の場合には外集団で、そのことを相手も知っている場合が共有知識条件ですね。相手が日本人なんだけど、相手はあなたが日本人だということを知らない、オーストラリア人かもしれないと思っているというのがプライベート知識条件です。



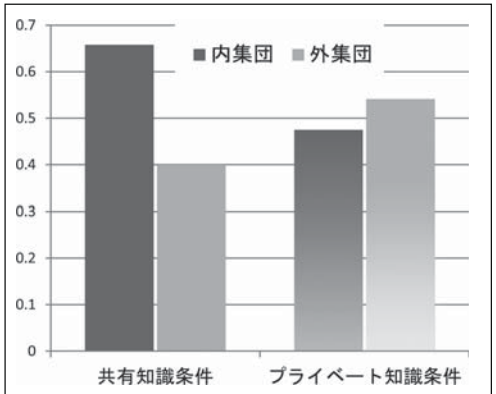
囚人のジレンマ実験での協力率。Yamagishi, T., Makimura, Y., Foddy, M., Matsuda, M., Kiyonari, T., & Platow, M. (2005)/ Comparisons of Australians and Japanese on group-based cooperation. *Asian Journal of Social Psychology*, 8, 173-190.

共有知識・外集団条件では、お互いに相手が違う国の人だということを知っています。日本人はここであんまり大きな差はないんですけれども、オーストラリア人は共有知識条件でだけ内集団を優遇するようになる。

これまでの実験は囚人のジレンマを使った実験なんですけれども、これをディクテーターゲーム、つまり独裁者ゲームでやっても同じような結果が生まれます。次の図がその結果です。

独裁者ゲームは、１人が一方的にお金を分けるゲームです。そのときに、共有知識条件とプライベート知識条件で、相手に分配してあげる金額を比較します。集団は最小条件集団です。この実験でも、共有知識条件、つま

りお互いに所属集団が分かっているときには、自分の集団を優遇するようなことをします。だけど自分だけが相手の集団を知っている場合ですと、そういう差をつけることをしないんですよね。むしろちょっと逆の傾向があります。



独裁者ゲームで、相手に分配する金額の比率。Yamagishi, T., & Mifune, N. (2008). Does shared group membership promote atruism? *Rationality and Society*, 20(1), 5–30.

次に紹介するのは独裁者ゲームじゃなくて順序付き囚人のジレンマゲームを使った実験です。あんまりオタク的な話をしてもしょうがないんですけれども、囚人のジレンマは、普通、２人が一斉に決めますよね。そうじゃなくて１人が決めて、その結果を相手に知らせて、その後で相手が決めるというのが、この順序付き囚人のジレンマゲームです。

面白いことに、このシーケンシャルPDゲームで最初に決定をするファーストプレーヤーは、相手が自分の集団の人か別の集団の人かで、相手に提供する金額に差をつけません。ところが、ファーストプレイヤーが協力の手を選択した後でセカンドプレイヤーが選択する場合、セカンドプレーヤーは相手が自分の集団の人かどうかで差をつけます。なぜファーストプレイヤーとセカンドプレイヤーでそんな違いが生まれるかというと、ファーストプレーヤーが協力しないのは、あんまり悪いことじゃないからです。

この実験とは別に、ファーストプレイヤーとセカンドプレイヤーが協力しなかった場合、それがどれぐらい悪いことですかと参加者に聞いているデータがありますけれども、そうしたデータでは、多くの人たちはファーストプレイヤーが協力しないのはあまり悪いことではないと思っています。ファーストプレイヤーは相手（セカンドプレイヤー）が協力してくれるかどうか分からない状態で自分は協力するかどうかを決めないといけません。自分が協力しても相手が協力してくれないかもしれない。だからファーストプレイヤーが協力しないのは、相手が協力してくれれば協力してもいいんだけど、自分だけ協力して馬鹿を見るのはいやだなと思って、協力しない可能性が大きい。心配になって協力しないのは、それはそんなに利己的な人間じゃないわけですね。そんなに悪いことじゃない。

だけどこの実験でのセカンドプレーヤーは、相手が協力したことが分かっているのに協力をしない。これはむちゃくちゃ悪いことですね。相手が協力してくれた。だけど俺は全部取っちゃうというのはよっぽどひどいやつです。つまり、非協力をしたという評判が回ったときに、その評判の悪さが全然違うんですね。ファーストプレーヤーとして協力しなかったのは、あんまり悪い評判につながらない。だからそのときには、自分の仲間に評判が伝わる可能性がある（ような気になる）場合でも、集団で差をつけるということはあんまりしません。

ところがセカンドプレーヤーの場合には、こんなことが分かったら大変な目に遭ってしまうと、非常に評判を気にするようになって、特に自分の集団の人間から同じ集団の仲間だとわかっているときに協力するんですね。

実はこれでもうそろそろ話を終わりにしたいと思いますけれども、同じような実験を、共有知識条件とプライベート知識条件で分けるんじゃなくて、目があるかどうかで分けるという実験をやってみました。これは先ほどお話しした独裁者実験では、プライベート知識条件では、相手が自分の集団の人間なのかどうかで、相手に分けてあげる金額に差をつけません。ところが同じプライベート知識条件でも、実験に参加した人が見ているコンピュータの画面に様式化した目の図を出しておくと、自分の集団の人にだけたくさん分けてあげるようになんです。

ということで、東京都で防犯のステッカーとして使っている歌舞伎役者の隈どりの絵を、一応許可を取って使わせてもらいました。そうしたら、やっぱり、目があると自分の集団に協力するようになるという結果が得られています。プライベート知識条件でも、自分の集団にたくさんお金を渡すようになる。ところが目がないと集団の差をつけない。

この実験の結果は、我々が今までやってきた知識の操作が、基本的にはレピュテーションが集団の中で回るということを操作しているのだということを示唆するものです。実際に集団の中で評判が回るわけじゃなくて

も、そういう状況なんだということを示すような手掛かりを何らかの形で与えると、それに応じた行動を取るようになるということですね。そういう心の仕組みを人間はすでに身に付けている。

　これはたぶん人間の社会性を考えるときに、考えておかないといけない人間性の一部ではないだろうかと考えています。どうも長い時間を取ってしまいまして申し訳ありませんでした。これで終わります。

● 参考文献：

(1) Toshio Yamagishi, “The Social Exchange Heuristic: A Psychological Mechanism That Makes a System of Generalized Exchange Self-Sustaining” In Mark Radford, Susumu Ohnuma, and Toshio Yamagishi (Eds.), Cultural and ecological foundations of the mind. Hokkaido University Press. pp.11-37(2007)

(2) Nobuhiro Mifune, Hirofumi Hashimoto, Toshio Yamagishi “Altruism toward in-group members as a reputation mechanism” Evolution and Human Behavior, 31(2), 109-117. (2010)