

価値・予測・誤差

—社会性を支える意思決定システム—

大平英樹 (名古屋大学)

Value, prediction, and error: Decision-making systems underlying social mind

Hideki Ohira (*Nagoya University*)

(2016年6月4日受稿, 2016年6月25日受理)

Though brain regions which are related to social behaviors and social affects have been widely elucidated, the mechanisms how those regions determine social mind are not understood. Here I suggest that a theoretical framework inspired by research on decision-making and reinforcement learning might be useful to consider the mechanisms underlying human's social mind. Specifically, it has been argued that three systems of decision-making, the Pavlovian system, goal-directed system, and habit system interact, based on evaluation of values, to determine wide ranges of behaviors of humans and animals. The present article proposes that such framework can be expanded to social behaviors and social affect, to shed new lights on interpretations of significance of the social phenomena and to draw new hypotheses for future empirical studies.

Key words: Social behavior, social affect, decision-making, reinforcement learning, value

1. はじめに

日本感情心理学会第23回大会は「社会性と感情」をテーマとして開催された。このテーマに関連する企画として「感情の脳科学」と題するシンポジウムが行われ、筆者は指定討論を務めた。本稿は、同シンポジウムを契機として、ヒトの社会性を実現する意思決定の諸システムと、その背後にある神経メカニズムについて論考するものである¹。また本稿は、筆者が「エモーション・スタディーズ」第1巻第1号に掲載した拙論(大平, 2015)と関連が深く、その続編として位置付けられる。

Correspondence concerning this article should be sent to: Hideki Ohira, Department of Psychology, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan (e-mail: ohira@lit.nagoya-u.ac.jp)

¹ ただし本稿は、シンポジウムでの講演や筆者による指定討論に立脚しつつも、必ずしもそれらを正確に報告するものではなく、新たに論考を展開したものである。

2. 社会性に関連する脳部位

我々は、他者に共感し、他者と意思疎通を図り、他者と協力したり、交渉したり、時には競ったり争ったりして、社会生活を営んでいる。そうした他者との相互作用の過程で、我々はさまざまな社会的感情 (social affect) を経験する²。今回のシンポジウムでは、近年の研究により社会性や社会的感情に関連する脳部位とその機能が解明されつつあることが報告された (Figure 1)。

罪責感 (guilt) は社会規範からの逸脱に伴って経験される感情であり、羞恥心 (embarrassment) は適切な振る舞いの失敗などにより経験される感情である。これらの負の社会的感情は、内側前頭前皮質 (medial prefrontal cortex: MPFC) と上側頭溝 (superior temporal sulcus: STS) の賦活を惹起する (Takahashi,

² 本稿では、情動 (emotion)、気分 (mood)、主観的感情体験 (feeling) などを包括する広範な概念として感情 (affect) の語を用いる。

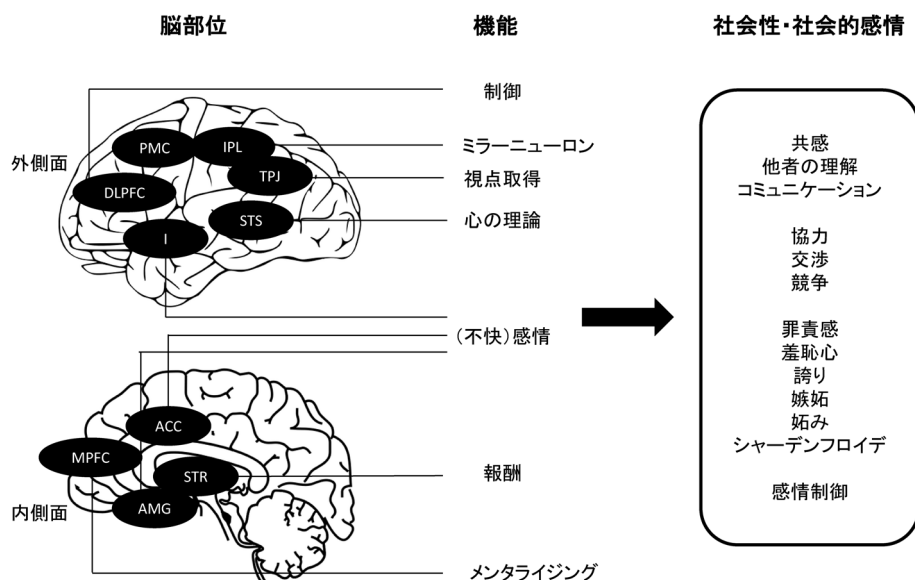


Figure 1. 社会性と社会的感情に関連する脳部位とその機能。DLPFC: 背外側前頭前皮質 (dorsolateral prefrontal cortex), I: 島 (insula), PMC: 前運動野 (premotor cortex), IPL: 下頭頂小葉 (inferior parietal lobule), TPJ: 側頭・頭頂結合 (temporal-parietal junction), STS: 上側頭溝 (superior temporal sulcus), MPFC: 内側前頭前皮質 (medial prefrontal cortex), ACC: 前部帯状皮質 (anterior cingulate cortex), STR: 線条体 (striatum), AMG: 扁桃体 (amygdala)

Yahata, Koeda, Matsuda, Asai, & Okubo, 2004)。MPFCとSTSはまた、何かを達成した喜びと他者からの賞賛に伴い経験される誇り (pride) の感情に伴っても活動する (Takahashi, Matsuura, Koeda, Yahata, Suhara, Kato, & Okubo, 2008)。また我々は、重要な他者との人間関係が第三者により脅かされる際に嫉妬 (jealousy) を経験する。嫉妬はSTSに加えて扁桃体 (amygdala) や島 (insula) の賦活を導く (Takahashi, Matsuura, Yahata, Koeda, Suhara, & Okubo, 2006)。これに対して、自己が他者より劣っているという比較が意識される時に湧き上がる感情が妬み (envy) である。妬みは、痛みの知覚に関連する前部帯状皮質 (anterior cingulate cortex: ACC) の背側部を活動させる (Takahashi, Kato, Matsuura, Mobbs, Suhara, & Okubo, 2009)。ところが、妬みの対象が失墜すれば、我々は密かな喜びを経験することがある。この感情はドイツ語でシャーデンフロイデ (Schadenfreude: 「恥知らずの喜び」) と呼ばれるが、その度合いは線条体 (striatum) の腹側部の活動強度により規定される (Takahashi et al., 2009)。これら多くの社会的感情を支える脳部位はまた、運動前野 (premotor cortex: PMC) ・下頭頂小葉 (inferior parietal lobule: IPL) ・島から成るミラーニューロン・システム³ (Ishida, Suzuki, & Grandi, 2015) と共に、共感 (empathy) の神経ネットワークをも構成している (Zaki & Ochsner, 2012; 大平, 2015)。さらに、他者への共感や

他者の心の理解は、背外側前頭前皮質 (dorsolateral prefrontal cortex: DLPFC) の働きにより調整される (Himichi & Nomura, 2015; Himichi, Fujita, & Nomura, 2015)。

これらの知見は、社会性や社会的感情が、扁桃体や島、ACC背側部などの痛みや不快感情に関連する脳部位、快感に関与する腹側線条体などの報酬系、MPFCを中心とするメンタライジング⁴ (mentalizing) のシステム、STSや側頭・頭頂結合 (temporal-parietal junction: TPJ) から成る心の理論⁵や他者視点取得のシステム、の働きに関連することを示唆している。そして、制御の座だと考えられているDLPFCが、個人が置かれた社会的文脈を考慮しつつ社会的感情を調整する機能に関連するのだと解釈できるであろう。

³ ヒトを含む霊長類の脳内で、自ら行動するとき、他の個体と同じ行動をするのを見ているときの、両方で活動する神経細胞をミラーニューロンと呼ぶ。マカクザルの前運動野で最初に発見されたが、ヒトでは前運動野と下頭頂葉を含む広範な脳領域で同様な働きが観測されるため、ミラーニューロン・システムと呼ばれている。ミラーニューロン・システムは、他者の行動の模倣、他者の感情や意図の理解、などの基盤となっていると考えられている。

⁴ 自己や他者の、意図や感情などの心的状態を推測し、統一的に表象する過程をメンタライジングと呼ぶ。

⁵ 他者の心の状態、目的、意図、知識、信念、志向、疑念、推測などを推測する認知的機能を心の理論と呼ぶ。

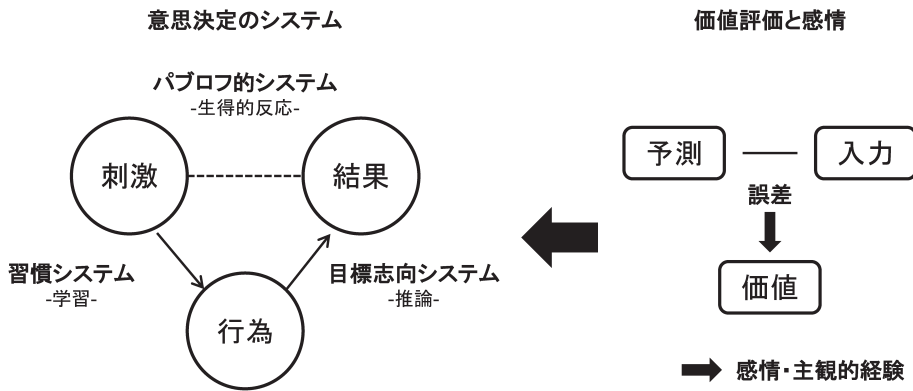


Figure 2. 意思決定と価値評価の原理 (Gesiarz & Crockett (2015) を基に筆者が改変)

社会性や社会的感情に関連する脳の構造や機能の解明は、確かに大きな学術的成果であった。しかしながら、こうした研究は本質的に、個々の心的機能や社会的感情に対応する脳部位を探るマッピングであり、上記の脳部位がどのようにして社会性や社会的感情を創発するのかについては語るものではない。実際、上記のように、比較的少数の脳部位の組み合わせが、社会性に関わる広範な心的現象に関連している。そうだとすれば、個々の脳部位の機能を、個々の心的現象や社会的感情に対応づけて解釈しようとするには無理がある。今後の研究においては、単なる脳マッピングではなく、ヒトの社会性がどのような原理により形成され、作動しているのかが問われるべきである。

そこで本稿では、そのような新たな探求の可能性を探るために、強化学習 (reinforcement learning) に基づく意思決定 (decision-making) の観点から、社会性の原理を考えたい (Quartz & Asp, 2015; Gesiarz & Crockett, 2015)。意思決定とは、刺激や状態の価値 (value) を評価し、それに基づいて行動を選択することであり、進化の過程で形成され多くの動物が共有する、生きていくための最も基礎的な仕組みである。それゆえ、意思決定の仕組みは、脳神経系を中心とした生物学的な基盤に実装されていると考えて不自然ではないであろう。以下では、ヒトにおいて特有であるように思える複雑な社会性が、こうした基礎的な生物学的枠組みにより、どのように実現されているのかを考える。

3. 強化学習に基づく意思決定システム

強化学習は、学習心理学の知見に触発されて主として工学の分野で発達した、ある主体 (agent) が自律的に学習し意思決定を行う過程を表現する理論的枠組みと計算論的アルゴリズムである。強化学習理論では、ヒト、動物、ロボットなどの多様な主体の、多様な場面における学習と意思決定を、刺激 (stimulus)、

行為 (action)、結果 (outcome) という3項の随伴性により簡潔に表現する (Figure 2)。主体の行動は、これら3項の随伴性の学習により規定されるが、その様態として、(1) パブロフ型システム (Pavlovian system)、(2) 目標志向システム (goal-directed system)、(3) 習慣システム (habit system) という3つの意思決定システムが考えられている (Balleine & O'Doherty, 2010; Dolan & Dayan, 2013)。

3.1 パブロフ型システムと価値

パブロフ型システムは、食物、水、適切な温度、異性、など生存や繁殖のために必要な事物を報酬とみなして正の価値を与える生得的な仕組みに基づいており、それらの報酬を予告する刺激に対する反射的な獲得行動の学習を促す。一方、痛み、毒物、捕食者、同性のライバルなど生存や繁殖に不利な事物には負の価値が付与され、それらを予告する刺激を回避する行動の学習が促進される。甘い味や匂いは栄養やカロリーの信号であり、未知の食物であってもこうした特徴を有していれば、空腹時には反射的に摂食行動を誘発する。ところが、それが昔よく使用された「猫いらず」のように実は毒物であった場合、それを摂食して運よく生き残ったネズミは同じような味や匂いを持つ食物を二度と食べようとはしない (味覚嫌悪条件付け)。これらは、心理学で古典的条件付け (classical conditioning)、またはパブロフ型条件付け (Pavlovian conditioning) と呼ばれる過程とほぼ同義である。強化学習理論では、この事態を次のように数理化する。ある主体が、時刻 t において、ある環境の状態 $s(t)$ のもとで (例えば特定の刺激が呈示されるなど)、報酬 $r(t)$ を得るとする。この時点で、この主体が、その状態のもとで将来得られると見込まれる報酬総額の期待値 ($V(s(t))$) は次式のような価値関数で表現される。

$$V(s(t))=E[r(t)+\delta r(t+1)+\delta^2 r(t+2)+\dots]$$

E は期待値であることを示し、 δ は割引率 (discounting rate: $0 < \delta < 1$) である。これは、遠い将来の報酬ほど、その価値が少なく見積られることを表している。この価値関数の時間差分を報酬予測誤差 (reward prediction error) と呼び、次式のように表現される。

$$\varepsilon(t)=r(t)+\delta V(s(t+1))-V(s(t))$$

これは、獲得した報酬の価値が、ひとつ前の時点で予測した価値からどれくらいずれているかを表している。主体は、このずれをゼロにすることを目指して学習を行う。Figure 2の右側はこの原理を表現している。

ここで重要なのは、強化学習理論では、感情と認知を区別しないことである。強化学習理論においては、価値の評価こそが感情であり、それは価値関数の認知的計算に基づいている。価値は意思決定システムにおける計算過程で生じている現象であるが、主体がヒトやある種の動物などのように意識を持っている場合には、価値の計算に伴って、喜び、落胆、嫌悪など我々が呼ぶような主観的経験が体験される。この観点においては、主観的経験は、価値の計算の過程や結果を言語の能力によりカテゴリー化し表象したものであり、それ自体は行動の原因でも結果でもない。もうひとつ重要なのは、この価値計算は身体的感覚に基づいていることである。ラットは、同じ甘い味であっても、カロリーのない人工甘味料よりもカロリーのある糖を好み、後者を予告する刺激を学習することができる (カロリー条件付け)。これはカロリー摂取に伴う血糖値の上昇や、それをモニターしたことによるある種の身体的感覚に起因するはずである。また、上記の毒餌による中毒により生じる味覚嫌悪条件付けは、腹痛や吐き気⁶などの身体的感覚により媒介されるに違いない。変温動物であるトカゲも、多くの餌が得られるのであれば、普段は嫌う低温環境に敢えて踏み込んでいく。この際の、報酬としての餌と低温環境への嫌悪は、線形なトレードオフの関数関係を示す (Cabanac, Cabanac, & Parent, 2009)。このような、本来、比較が不可能なはずである複数の事象を同一の尺度に射影して比較し選択できるようにするための、いわば脳内の共通貨幣 (common currency) としての仕組みが価値なのであり、その背後にあるのが中脳-線条体のドーパミン神経系、いわゆる脳内の報酬系の機能である。そして、その時に報酬系における価値計算が表象されたもの、さらにはその一部が意識化され体験され

たものが、感情である。この際に、報酬系に価値を計算する材料を提供するのは、扁桃体や島の機能に裏付けられた身体的な感覚に他ならない。すなわち、強化学習理論の観点においては、感情とは「身体化された価値」⁷を意味する。

3.2 目標志向システムと習慣システム

一方、心理学で道具的条件付け (instrumental conditioning)、あるいはオペラント条件付け (operant conditioning) と呼ばれる過程を実現するのが目標志向システムと習慣システムである。目標志向システムは、現在達成すべき複数の目標をリスト化したうえで、その優先順位を付ける (目標の価値の評価)。その上で、行為とその結果の随伴性を推論し、最も価値の高い目標に近づく最適な行為を選択する。これを実現するため、このシステムは意識的で熟慮的な計算過程に支えられており、柔軟性が高く状況の変化に敏感に対処することができるが、その分認知負荷が高い。それゆえに、目標志向システムの効率は利用可能なワーキングメモリ容量に依存する。この特性から想像できるように、目標志向システムではDLPFCが最も重要であり、そこに価値を評価する線条体や、価値を行動に変換するACCが協調することで機能する (Kaller, Rahm, Spreer, Weiller, & Unterrainer, 2011; Balleine & O'Doherty, 2010)。これに対して習慣システムは、一回ごとの行為の結果より、過去の経験に基づいた行為の価値を評価する。過去において多くの報酬をもたらした行為には、現在の結果がどうであろうとも、高い価値が与えられる。いったん習慣化した行為は、その手掛りが与えられれば、半ば自動的に発動されるので、認知的負荷は低くワーキングメモリ容量にも依存しない。その反面、習慣システムは融通が利かず、状況の変化に機敏に追随することは困難である。習慣システムを支える計算は、腹側線条体と背側線条体により担われている (O'Doherty, Dayan, Schultz, Deichmann, Friston, & Dolan, 2004; Schultz, 1998)。その原理は、生得的に実装されているパプロフ型システムの価値評価の仕組みを利用して、刺激に対する行為の価値を、予測と結果の報酬予測誤差により評価し、その価値の重みづけにより確率的に行為を選択するという形で数理化される。

目標志向システムと習慣システムは、時と場合により、拮抗的にも補完的にも働く。体重増加に悩んでいる人物がスーパーマーケットで食材を買う場面を想像しよう。この人物は、健康を維持する、食欲を満たす、嗜好を満足させる、など多様な目標を持っている。その中でどれを優先するか、そのためにどのような購買

⁶ しかし、ラットは嘔吐する生理的機能を有していない。

⁷ この考え方は、Prinz (2004) の感情の哲学理論と関連が深い。Prinz (2004) の理論については、戸田山 (2016) を参照。

行動をすべきかを決めるのは目標志向システムである。一方、習慣システムは、過去に身体的な快楽と結びついた報酬をもたらした油っぽく高カロリーな食材を買うことを動機付ける。この段階では、両システムの働きは明らかに拮抗的である。しかし、いったん健康維持を目標として選択したならば、目標志向システムと習慣システムの拮抗状態は解消され、具体的にどの商品を選択するかはほとんど習慣システムに委ねられる。また、身体的快楽を優先することを選択した場合も、やはり具体的な商品選択は習慣システムに依存して行われる。スーパーマーケットの中は、熟慮的には処理しきれない程の情報に溢れており、「これまでいつも買っていた」ことを選択の基準とせざるを得ないからである。このため、初めて訪れたスーパーマーケットでは、どの商品がどこに陳列されているかが未知なので習慣を利用することができず、購買行動に困難を感じるようになる。このように、我々の意思決定の多くの部分は、実は習慣システムにより規定されていると考えることができる。

4. 意思決定システムの社会性への拡張

4.1 社会的価値

我々の社会的行動も、こうした意思決定システムが司っていると考えることができる (Geislarz & Crockett, 2015)。進化心理学では、社会における適応度を高めることが我々の目標であり、その実現のために重要なのは「評判」(reputation) を高めることであると考えられている (Nowak & Sigmund, 2005)。評判が高くなければ他者にゲームのパートナーとして選択してもらえず、他者からの協力を得て利益を獲得することができない。それは社会環境における適応度を著しく低めてしまう。そこで、自己の評判高揚を実現する事象が報酬とみなされ、正の社会的価値が与えられる。逆に自己の評判を下げる事象には負の社会的価値が与えられる。

社会的価値の計算は、食物、水、痛み、捕食者などに対するのと同じように、線条体において作動するパブロフ型システムにより実行されると考えられる。実際、他者からの賞賛によりもたらされる社会的価値の高揚は、金銭的報酬を得た時と同じように、線条体を賦活させる (Izuma, Saito, & Sadato, 2010)。この時、空腹の際に食物にありついた時と同じように、線条体で計算される価値は、快感情として体験されるのであろう。この意味において、社会的価値の高揚に伴う快感情は、社会における自己の評判、すなわち社会において自己がどの程度適応し成功できるかという予測の指標として機能する。一方、他者からの低い評価や排除、不平等な扱いなどは、強い不快感情を惹起する。しかも、こうした社会的価値の低減に伴う不快感情は

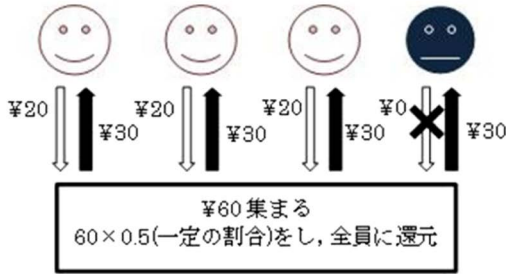
身体的反応を伴う。サイバー・ボール課題⁸ (cyber ball paradigm) で実験的に仲間外れを経験させられた個人の身体には炎症反応が生じる (Slavich, Way, Eisenberger, & Taylor, 2010)。最後通牒ゲーム⁹ (ultimatum game) で不公正な提案を受けた個人は、自律神経系の覚醒反応を示す (Osumi & Ohira, 2009)。つまり、社会的場面で我々が経験する快や不快の感情は、その時点における我々の社会的価値を符号化しているものであり、その社会的価値もやはり、食物や痛みなどの場合と同じように、身体化されている。社会心理学では、こうした感情により尺度化された社会的価値をソシオメーター (sociometer) と呼ぶ (Leary & Baumeister, 2000)。これは、自己が社会的に承認されている程度を認識するものであり、その「測定値」を我々は、自尊感情 (self-esteem)、誇り、羞恥心、罪責感などと呼ばれる社会的感情として体験する (Quartz & Asp, 2015)。

こうした社会的価値の評価のためには、他者が自己をどう認識するかを推測し、他者が見た自己の姿を想像して表象する能力が必須である。ここで重要なのは特定の他者から良く思われることではなく、社会の中の自己の評判を高めることであるから、社会的価値の評価には特定の他者を超えて内在化された一般的な他者の視点を表象できる能力が必要とされる。我々は、現実には存在しない、特定の誰でもない他者から見て、自分自身がどれ程、優れた人格者であるか、偉いか、カッコいいか、などのように自己の社会的価値を評価するのである。このように考えれば、上述したように、多くの社会的感情の経験に伴って、MPFC、STS、TPJなどの、メンタライジング、心の理論、他者視点取得に関わる神経ネットワークに加えて、パブロフ型システムの中核である線条体の報酬系が働くことは容易に理解できる。それゆえに、心の理論の機能に障害のある自閉スペクトラム症候群の患者は、

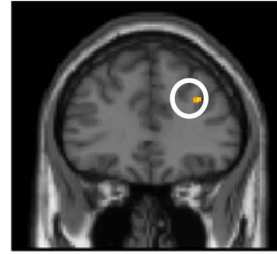
⁸ 実験参加者が、2人の他者と、コンピュータ上で自由にキャッチボールをする課題。途中から、他の2人からボールを投げてもらえなくなることで、排除や孤独の感覚を与えることを目的としている。極めてシンプルな課題であるが、被排除感や孤独感を頑健に惹起できるとされている。

⁹ 2人のプレイヤーが一定の金額を分配する、交渉事態をシミュレーションに表現する経済ゲーム。提案者 (proposer) と呼ばれる一方のプレイヤーが一方的に配分の方針を提案でき、応答者 (responder) と呼ばれる他方のプレイヤーは、その提案を受諾するか拒否するかを決定する。受諾すれば提案どおりに金額が配分され、拒否すれば両者とも金銭を受け取ることはできない。新古典主義的な経済学理論から考えれば、応答者はどれ程少ない配分であっても拒否した場合の0円よりはましなので必ず受諾し、提案者はそれを知っているので、自分に最大限の金額を与える。しかし実際には、あまりに不平等な提案はしばしば拒否され、提案者も極端に搾取的な提案はせず応答者に何割かを与える決定をすることが普通である。こうした行動は、経済学的な人間観とは反するために注目されるようになった。

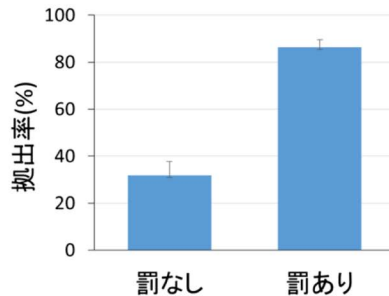
A. 公共財ゲーム



C. 背外側前頭前皮質



B. 公共財への拠出率



D. 島

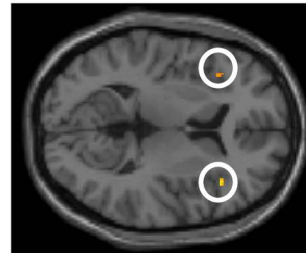


Figure 3. 公共財ゲームを用いた協力の認知神経科学的研究。A: 公共財ゲームの手続き。各プレイヤーは各試行において20円を拠出するか否かの意思決定を行う。この例では、最も右のプレイヤーは拠出していない。この結果、合計拠出額60円の50%である30円を全員が得る。拠出したプレイヤーの利益は差引10円であるのに対して、拠出しなかったただ乗り者は30円を得ることができる。B: ただ乗りへの罰がない条件と罰のある条件における公共財への拠出率。C: 公共財ゲームの罰のある条件において賦活する背外側前頭前皮質 (dorsolateral prefrontal cortex: DLPFC)。D: 公共財ゲームの罰のある条件において賦活する島

自己の評判に無頓着なのである (Izuma, Matsumoto, Camerer, & Adolphs, 2011)。そしておそらく、社会的価値の評価においても、予測誤差の原理が重要な役割を果たす。我々は、ある場面での将来の自己の行動が、他者の目にどう映るかを予測する。そして行動をした後の他者からの評価を想像して表象し、予測との比較を行う。我々は、この際の予測誤差が正の値であれば、自尊感情や誇りを体験し、予測誤差が負の値であれば、落胆や羞恥心を体験する。そうした社会的感情は、将来の行動を調整するために利用される。

このように、パブロフ型システムは、何であれ対象が表象できれば、その価値を計算することができる。そうであれば、我々が社会的価値を計算し、社会的意思決定を行うことができるのは、MPFC、STS、TPJなどの神経ネットワークにより、自己や他者の心的状態を対象として表象できる能力を持ったためであると考えることができる。

4.2 目標志向システムと習慣システムによる社会的行動のコントロール

社会的場面においては、複数の目標が葛藤すること

が少なくない。その一例としては、自己の利益を犠牲にしても公共の利益を優先するか、利己的に自己の利益を優先するか、という場面が考えられる¹⁰。目標志向システムは、ちょうどスーパーマーケットにおける買い物に際して健康維持と嗜好満足の目標間葛藤を調整するように、そうした社会的目標間の葛藤をも調整し得ると推論することができる。

筆者らは、この問題を検討するために、公共財ゲーム (public goods game) と呼ばれる経済ゲームを用いた研究を行った¹¹。このゲームでは、各プレイヤーは、集団共有のプールに資金を拠出するかどうかの意思決定を行う。集団全体の合計拠出額が運用された結果として、合計拠出金額に一定割合をかけた金額

¹⁰ ここでいう「自己」は、必ずしも個人のことでない。個人としての自己が同一視する集団、企業、地域、国家、などに拡張して考えることもできる。昨今噴出している企業の不正や、地域や国家のエゴなどの事例を考えれば、これが普遍的な問題であることが理解できる。

¹¹ この研究は、平成24-27年度に、科学研究費補助金・基盤研究 (A) 『正義観の心理・神経・生理・進化的基盤』 (代表研究者：大平英樹、研究課題・領域番号24243068) により行われた。

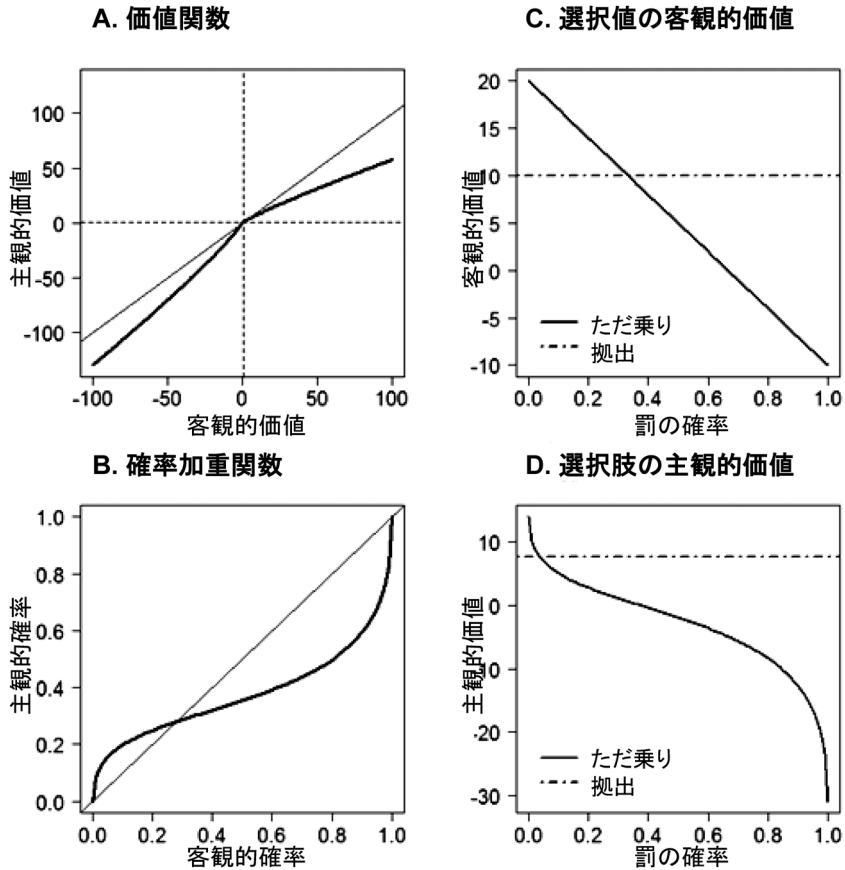


Figure 4. プロスペクト理論により説明される公共財ゲームにおける価値の評価。A: プロスペクト理論が主張する主観的価値を表す価値関数。細線は客観的な価値を表す。太線で表現される主観的価値は、正の価値も負の価値も過減性を示す。B: プロスペクト理論が主張する主観的な確率の知覚を表す確率加重関数。細線は客観的な確率を表す。太線で表現される主観的な確率の知覚では、稀な事象は実際より確率が高く感じられる。C: 罰の確率により規定されるただ乗り（実線）と拠出（協力：一点鎖線）の客観的価値。協力する場合には罰の影響はない。ただ乗りは、罰の確率が30%を超えると損になる。D: 罰の確率により規定されたただ乗り（実線）と拠出（協力：一点鎖線）の主観的価値。協力する場合には罰の影響はない。ただ乗りは、罰が10%程度の低確率であっても損に感じられる

が各個人に均等に利益として与えられる。我々の研究では、多数回の試行を可能にするため、1試行での拠出金は20円、運用率は50%とした。このゲームでは、全員が協力し拠出すれば、全員が獲得できる利益が大きくなる。ところが、資金を拠出しないただ乗り者は、コストを支払うことなく利益だけを獲得できる。しかし、全員がただ乗りを行うと合計拠出額は0円となり、誰も利益を得ることができない (Figure 3A)。このゲームでは一般に、ただ乗りして自己利益をはかる利己的動機の誘因が強いために、試行を重ねるうちに協力（資源の拠出）は減少する（「共有地の悲劇」）。この状況を打破し協力を増加させる仕組みとして、ただ乗りに対する罰が挙げられる (Fehr & Gächter, 2000)。そこで我々の研究では、罰がない条件と、ただ乗りには10%の確率で30円の罰金が科さ

れる条件を設定した。

Figure 3Bに示すように、罰の導入により拠出率は劇的に増加した。ところが、罰の確率は10%に過ぎないため、罰があってもただ乗りを続けた方が獲得できる金額の期待値は大きくなる (Figure 4C)。それに関わらず、この条件で拠出率が極めて大きいという結果は、個人が罰の恐れや不安などの感情に影響されて不合理な意思決定を行ったためであると解釈されるかもしれない。この問題について我々は、行動経済学において有名なプロスペクト理論 (prospect theory: Kahneman & Tversky, 1979) に基づく説明を試みた¹²。この理論によると、ある選択枝の価値は価値と

¹² プロスペクト理論に基づくシミュレーションは、共同研究者の片平健太郎（名古屋大学）が行った。記して感謝する。

確率荷重の積で表される。価値関数は客観的な効用の参照点からの差分 x の関数として次の式で与えられる (Figure 4A)。

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\beta & \text{if } x < 0 \end{cases}$$

ここで各パラメータは、この理論において標準的な値である、 $\alpha=0.88$, $\beta=0.88$, $\lambda=2.25$ とした。確率荷重関数は客観的確率の関数として次の式で与えられる (Figure 4B)。ここでもパラメータは標準的な値である、 $\gamma=0.5$ とした。

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$$

これらの条件を前提として、個人において計算される主観的価値を推定すると Figure 4D のようになる。つまり、ただ乗りにより7%以上の確率で30円の罰金が科されるならば、協力する方が期待される主観的価値は高くなるのである。こうした知見から、我々の研究の実験参加者は、単に感情に影響されて不合理な決定をしたのではなく、選択肢の価値をよく勘案して主観的ではあるが合理的な社会的意思決定を行っていたのだと解釈できる。実際に、罰のある条件では、罰のない条件に比べて、目標志向システムの中核である DLPFC が強く活動していた (Figure 3C)。また、罰のある条件では身体からの信号が表象される島の前部にも強い活動が観測された (Figure 3D)。この結果は、社会的価値もやはり、身体化された評価として符号化されていることが示唆される¹³。

さらに、同じ公共財ゲームを用いた別の研究で我々は、プレイヤー相互による罰と、法律のアナロジーである第三者的な権威による罰を操作した実験を行った。プレイヤー相互による罰の条件では、各プレイヤーは20円のコストを支払えば自由に、他のプレイヤーに30円の罰金を科すことができる。一方権威による罰の条件では、ただ乗りに対する30円の罰金は、10%の確率で均等に下される。両条件において、公共財への資金の拠出率はほとんど同じであった。しかし、プレイヤー相互による罰の条件では、1試行ごとに、拠出するかただ乗りするか意思決定の変動が大きかったのに対して、権威による罰の条件では拠出の意思決定は個人内では安定していた。プレイヤー相互

による罰の条件では、自分自身の利益を最大化するには、他者が協力するかただ乗りするか、他者がただ乗りを罰するか否か、それに基づいて自分自身がただ乗り者を罰するか否か、などの極めて複雑な駆け引きが必要となる。これはまさに目標志向システムが関与する過程であろう。これに対して、権威による罰の条件では、いったん選択肢間の社会的価値が算出されたならば、それに基づいて習慣システムに依存した意思決定をすればよい。法律のような制度や正義を重んじる社会的規範は、いったんそれが成立したならば、後はそのルールに従えばよいという習慣的システムの働きによって個人の行動を規定しているのかもしれない。

5. 結論に代えて

—社会性を意思決定の視点から考えることの利点—

最後に、本稿で提案したような強化学習に基づく意思決定の視点から社会性を考えることの利点を指摘しておきたい。

第一に、強化学習はヒトや動物の意思決定をシンプルに数理化するために考案された理論的枠組みである。それゆえに、これを適用することにより、一見複雑なヒトの社会的行動や社会的感情の背後にあるメカニズムを、統一的かつ定量的に表現し検討することが可能になり、ヒト以外の動物との比較検討も可能になる。第二に、こうした理論的枠組みを適用することによって、実証的に検討することが可能な研究仮説を多く導出できる。例えば、社会的意思決定の場面において、目標志向システムの働きが賦活されたならば、より合理的な選択がなされると予測される。実際、目標志向システムの中核である DLPFC を経頭蓋直流電気刺激 (transcranial direct current stimulation: tDCS) により賦活すると、交渉が必要な最後通牒ゲームでは平等な金銭分配が促進されるが、一方的に分配方針を強制できる独裁者ゲーム¹⁴ (dictator game) では利己的な分配が優勢になることが示されている (Ruff, Ugazio, & Fehr, 2013)。そして最も重要なことは、強化学習に基づく意思決定理論の枠組みに立脚すれば、どのように複雑な対象であろうとも、その表象が形成できれば価値を評価することができ、その価値に基づく意思決定は、食物、水、痛み、捕食者などの基礎的な対象への意思決定と同じメカニズムで記述し理解することができるということである。このような視点から見れば、ヒトの社会性が成立する第一の要件は、さまざまな社会的場面での選択肢を分節化し、情報処理

¹³ この時、個人は、罰への恐れや不安、規範を遵守することへの義務感、公共の利益を優先することへの誇りや自尊心など、さまざまな社会的感情を体験するであろう。しかし、強化学習に基づく意思決定理論の観点からは、そうした個人内の主観的体験を問う必要はない。これがこの視点の利点であり、客観的に記述される意思決定システムから、複雑な社会性を統一的に説明できる可能性が生じる。

¹⁴ 最後通牒ゲームと似ているが、提案者のプレイヤーが自由に金銭の分配方針を決めることができ、他方のプレイヤーはそれを受諾するしかない、というゲーム。新古典主義的な経済学理論の観点からは、このゲームでは提案者は合理的にすべての金銭を独り占めすると予測される。

の単位として扱う能力であると思われる。社会的場面は極めて大きな情報量を含むと考えられるが、ヒトは、概念間のネットワークを形成する能力や、刺激等価性¹⁵ (stimulus equivalence) を成立させる能力により、そうした大きな情報単位を扱うことができるようになったと考えられる。そして社会性が成立する第二の要件は、本稿で論じたように、そうした社会的場面を表す情報単位を、他者の視点から眺めて表象化する能力であると考えられる。

意思決定理論に基づくヒトの社会性の研究は端緒に付いたばかりであり、本稿で論じた内容には多くの仮説や推測が含まれている。今後、こうした視点からの実証的研究が発展し、ヒトという種についての理解がより豊穡なることを期待する。

引用文献

- Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, **35**, 48–69.
- Cabanac, M., Cabanac, A. J., & Parent, A. (2009). The emergence of consciousness in phylogeny. *Behavioral Brain Research*, **198**, 267–272.
- Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, **80**, 312–325.
- Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The Journal of Economic Perspectives*, **14**, 159–181.
- Gęsiarz, F., & Crockett, M. J. (2015). Goal-directed, habitual and Pavlovian prosocial behavior. *Frontiers in Behavioral Neuroscience*, **9**, 135.
- Himichi, T., Fujita, H., & Nomura, M. (2015). Negative emotions impact lateral prefrontal cortex activation during theory of mind: An fNIRS study. *Social Neuroscience*, **10**, 605–615.
- Himichi, T., & Nomura, M. (2015). Modulation of empathy in the left ventrolateral prefrontal cortex facilitates altruistic behavior: An fNIRS study. *Journal of Integrative Neuroscience*, **14**, 207–222.
- Ishida, H., Suzuki, K., & Grandi, L. C. (2015). Predictive coding accounts of shared representations in parieto-insular networks. *Neuropsychologia*, **70**, 442–454.
- Izuma, K., Matsumoto, K., Camerer, C. F., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 17302–17307.
- Izuma, K., Saito, D. N., & Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*, **22**, 621–631.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263–292.
- Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., & Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cerebral Cortex*, **21**, 307–317.
- Leary, M. R., & Baumeister, R. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 32. San Diego, CA, Academic Press. pp. 1–62.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, **437**, 1291–1298.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**, 452–454.
- Osumi, T., & Ohira, H. (2009). Cardiac responses predict decisions: An investigation of the relation between orienting response and decisions in the ultimatum game. *International Journal of Psychophysiology*, **74**, 74–79.
- Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.
- Quartz, S., & Asp, A. (2015). *Cool: How the brain's hidden quest for cool drives our economy and shapes our world*. New York: Farrer, Straus and Giroux.
- 大平英樹 (2015). 共感を創発する原理 エモーション・スタディーズ, **1**, 56–62.
- Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, **342**, 482–484.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, **80**, 1–27.
- Slavich, G. M., Way, B. M., Eisenberger, N. I., & Taylor, S. E. (2010). Neural sensitivity to social rejection is associated with inflammatory responses to social stress. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 14817–14822.
- Takahashi, H., Kato, M., Matsuura, M., Mobbs, D., Suhara, T., & Okubo, Y. (2009). When your gain is my pain and your pain is my gain: Neural correlates of envy and Schadenfreude. *Science*, **323**, 937–939.
- Takahashi, H., Matsuura, M., Koeda, M., Yahata, N., Suhara, T., Kato, M., & Okubo, Y. (2008). Brain activations during judgments of positive self-conscious emotion and positive basic emotion: Pride and joy. *Cerebral Cortex*, **18**, 898–903.

¹⁵ 複数の刺激の間の機能的な交換可能性を意味する。例えば、子どもが母親から「イヌはどれ？」と尋ねられて犬の絵を選ぶことができ、同時に母親から「これは何？」と尋ねられて子どもが「イヌ」と答えられた時、「イヌ」という音声と犬の絵の間に刺激等価性が成立したという。ヒトでは他の動物種と比べて特に刺激等価性が成立しやすいことが示されており、刺激等価性は言語や高度な表象機能が発現するための基盤となる認知特性であると考えられている。

- Takahashi, H., Matsuura, M., Yahata, N., Koeda, M., Suhara, T., & Okubo, Y. (2006). Men and women show distinct brain activations during imagery of sexual and emotional infidelity. *NeuroImage*, **32**, 1299–1307.
- Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: An fMRI study. *NeuroImage*, **23**, 967–974.
- 戸田山和久 (2016). 恐怖の哲学——ホラーで人間を読む—— NHK 出版
- Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience*, **15**, 675–680.