

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 27 日現在

機関番号：32639

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26750379

研究課題名(和文)社会的行動学習における大脳基底核・ドーパミン系の役割

研究課題名(英文)Role of the basal ganglia and dopamine neurons for reward learning in social context

研究代表者

榎本 一紀(ENOMOTO, Kazuki)

玉川大学・脳科学研究所・科研費研究員

研究者番号：10585904

交付決定額(研究期間全体):(直接経費) 2,100,000円

研究成果の概要(和文):社会的状況では長期的な報酬期待に基づく意思決定が必要となる。ニホンザルを被験体として、複数の報酬獲得ステップを経てゴールに到達する行動課題を学習させ、中脳ドーパミン細胞が表現する将来報酬情報は、学習が過度に進むと超長期的な報酬情報を反映することを明らかにした。また、複数頭のサルを用いた社会的意思決定課題において、動物は他者の選択行動や報酬摂取行動を観察し、自身の学習に利用していることが示唆された。これらの結果は、社会のなかで適切な意思決定を行い、効率的に利益を獲得するための学習に関わる脳神経メカニズムの解明に貢献すると期待される。

研究成果の概要(英文): In a social context, we make decisions and select actions based on long-term reward expectation. We showed that activities of midbrain dopamine neurons in monkeys represented super-long-term reward information when monkeys were over-trained about a behavioral task in which they attain a goal following multiple reward acquisition. We also trained monkeys to do a social task in which they need to observe the other agent to obtain reward efficiently. Monkeys' behavior indicated that they learn from others' action and reward consumption to improve their own decisions. These results contribute to uncover the neural system involved in learning and decision-making in a social context.

研究分野: 基盤・社会脳科学

キーワード: ドーパミン 社会性 霊長類 報酬 意思決定 学習

1. 研究開始当初の背景

われわれ人間や多くの動物は相互依存的な社会関係のなかで生きている。複数の行動主体が相互依存的な関係をもつ状況では、他者の行動が自分の利益につながる場合があり、数ステップ後の行動結果を含んだ、長期的な報酬予測が必要となる。また、社会のなかで適切な意思決定・行動選択を行い、効率的に利益を獲得するためには、信頼できる他者の行動を観察し、必要な情報だけを抽出し、自己の行動に生かすことが必要である。そのような社会的環境における学習や意志決定・行動に関わる脳神経メカニズムの解明は、神経科学のみならず、社会学や精神医学の分野からも期待されている。最近の研究によって、他者の行動そのものや、それに基づいた学習に関わる大脳皮質の神経活動が報告されているが (Azzi et al, 2012; Hosokawa and Watanabe, 2012; Chang et al, 2013; Suzuki et al, 2012)、ヒト以外にも多くの動物に共通して学習や行動選択に関わるとされる大脳皮質下の、特に大脳基底核・ドーパミン系における神経活動が、他者の行動や報酬獲得によってどのように調節されるのかは未だ明らかになっておらず、より基盤的な神経メカニズムの理解には至っていない。

中脳ドーパミン細胞は、大脳皮質や辺縁系、線条体などの広範囲にわたる脳領域に投射し、それぞれの領域におけるシグナル伝達の可塑性を調節している。現在までの研究によって、ドーパミン細胞は様々な外界の事象に応答し、その放電頻度の変化は、刺激の新規性や報酬の期待値を反映することが知られている (Salamon et al, 2005)。報酬が条件刺激によって予告されると、条件刺激に対して報酬価値 (報酬量や報酬確率) を反映した応答を示し (Schultz et al, 1997; Morris et al, 2004)、実際に報酬が得られた/得られなかったときには、予測との誤差を反映した活動を示す (Sato et al, 2003)。また、ドー

パミン細胞から強い投射を受けている大脳基底核の線条体は、大脳皮質前頭前野や前頭内側皮質、運動野などの認知機能・運動機能を司る脳領域からも情報を受け取っている (Fudge, et al, 2002)。線条体細胞は報酬情報 (Hikosaka et al, 1989; Apicella et al, 1992) に加えて運動情報 (Alexander and Crutcher, 1990) もコードしていることが報告されており、行動の価値を表現することで (Samejima et al, 2005)、選択できる行動のなかで価値の高いものを選び、利得を増やす学習に関わる神経システムにおいて中心的な役割を担っていると考えられている。

研究代表者は現在までの研究において、ニホンザルのドーパミン細胞および線条体細胞の活動が、長期的な将来報酬の価値を表現することを報告した (Enomoto et al, 2011; Yamada et al, 2013)。それらの研究では、動物に意思決定を伴う行動選択課題を学習させ、中脳ドーパミン細胞および線条体細胞の活動を記録した。その結果、1 ブロック内で複数回 (2-3 回) の報酬が得られる課題の構造に習熟すると、神経細胞活動は複数の将来報酬情報を積算して表現するように変化することを証明した。しかしながら、より学習が進んだ時にドーパミン細胞が更に長期的な報酬情報を表現するのかどうかについては明らかではない。

また、これらの神経細胞の活動が、他者によって自己の学習効率や行動選択・報酬獲得が左右される社会的な状況においてどれほど影響を受けるのかについては報告が少ない。ヒトを対象としたイメージング研究によって、線条体は寄付行動 (Harbaugh, et al, 2007) や不平等性の検知 (Fliessbach et al, 2007) などに関わることが示唆されているが、一細胞単位での情報表現についてはほとんど知られておらず、大脳皮質下の脳領域も含めた神経回路の理解が進んでいない。よって、社会的環境における大脳基底核・ドーパミン

細胞の神経活動を明らかにすることは、多くの動物に共通した社会性に関わる脳神経ネットワーク解明のために必須である。

2. 研究の目的

(1) ドーパミン細胞の活動が長期的な報酬情報を表現するのであれば、目前の報酬だけでなく、数手先の試行でも報酬獲得が予測できる状況において、細胞活動に累積的な報酬情報が反映されると考えられる。ドーパミン細胞がどれほど長期的な将来報酬価値を表現するのかを検証するため、ニホンザルに複数ステップの報酬獲得を経てゴールに到達する行動課題を学習させ、課題遂行中の細胞活動を記録する。その後、強化学習モデルを用いた計算機シミュレーションと細胞活動を比較することで、計算理論に基づいた定量的な解析を行う。

また、近年の研究によって、ドーパミン細胞や線条体細胞の活動は部位によって応答特性が異なることが報告されているので (Matsumoto and Hikosaka, 2009) それぞれの細胞活動と記録部位について詳細な組織学的検討を行う。ドーパミン細胞については、特に背外側に位置する黒質緻密部と腹内側に位置する腹側被蓋野との違いに注目する。

(2) 複数のニホンザルを被験体とした社会的状況における行動実験を行う。ニホンザルやアカゲザルなどのマカクザルはヒエラルキー構造のある母系社会を営み、野生環境での観察において、毛繕い行動やエサの分配などの様々な社会的行動が報告されており (今西, 1957; 伊谷, 1951; Yamagiwa and Hill, 1998) 二者間の優劣やヒエラルキーの順位を推定するための指標について豊富な参照例がある。実験室内環境においては、アカゲザルは報酬に基づいた自己と他者の相互関係がある場合に、他者の報酬条件によって自

分の行動を変化させることが明らかになっている (Subiaul et al, 2004; Chang et al, 2011)。そこで、複数頭のニホンザルを用いて、両者が協調して行動することで効率的に報酬を得られる行動課題を学習させ、相手の行動を観察学習することで報酬を多く獲得できる関係を築く。行動課題は、向かい合った二頭が交互に行動選択を行うもので、効率的に学習して多くの報酬を得るために、相手の行動とその結果の情報が有用な場合は、それを観察し、自己の行動に生かすことが必要となる。その上で、相手の行動の信頼性をコントロールした条件を課題に組み入れることで、条件によって効率の良い観察学習の仕方を変化させる。行動パターンや反応時間、眼球運動などの行動データを詳細に検討することによって、計算理論に基づいた二者間の学習モデルを確立する。

3. 研究の方法

(1) ニホンザルを被験体として、1 ブロックが3回の報酬獲得試行からなる行動課題 (図1) を学習させ、課題遂行時のドーパミン細胞活動を電極記録する。

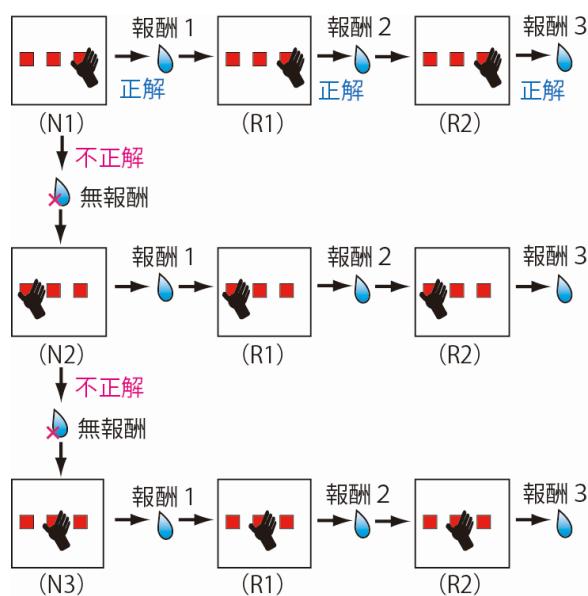


図1 行動課題1ブロックの流れ

この課題では、3つの選択肢(押しボタン)のうち1つが報酬を得られる正解ボタンとなっており、動物は試行錯誤で正解ボタンを探す。正解ボタンを選んだ時は、報酬として口元のパイプからジュースが与えられる。最初の試行(第1探索試行:N1)で正解ボタン以外を選び、報酬が得られなかった場合は、次の試行(第2探索試行:N2)で前試行とは別のボタンを選び、そこでも正解できなければ、更に次の試行(第3探索試行:N3)で、残った1つのボタンを選ぶことになる。いずれの場合でも1度正解すれば、続く2回の試行でも同じボタンを選択することで、2回目、3回目の報酬が得られる(第1繰り返し試行:R1、第2繰り返し試行:R2)。N1からN3にかけて報酬確率(報酬獲得率)は段階的に上がってゆき、R1・R2においてはほぼ確実に報酬が得られる。よって、もし動物が目前の報酬のことしか考えていないのであれば、各試行開始時に予測される報酬の価値は試行ごとの報酬確率と相関するが(将来報酬割引率: $\gamma = 0$)、1ブロック全体の報酬を見込んでいるのであれば、N1~N3開始時から、R1・R2の報酬も含んだ3回分の報酬を予測し($0 < \gamma < 1$)、更に先までの報酬予測をしているのであれば、次のブロックの報酬も含んだ、4回分以上の報酬情報表現が見られるはずである(1)。ドーパミン細胞は課題開始の合図である視覚刺激と強化因子に応答し、それぞれの応答は報酬価値や予測誤差を表現することが分かっているので、強化学習則に基づいた計算機シミュレーションによって算出した価値と実際の応答を比較することで、将来報酬の割引率などの定量的な解析を行う。実験終了後、脳標本を作製し、細胞の応答特性と記録部位を検討することで神経回路レベルの脳機能に迫る。

(2) ニホンザルはメスを中心とした階層的な社会構造を持つことが分かっているので、実

験には3頭のメスザル(体重約7kg)を用いる。飼育ケージは仕切りのある集団ケージを用い、ケージ内での親愛行動や威嚇行動、毛繕いなどの行動観察に加えて、実験室でのフードテストによってヒエラルキーの順位を推定する。動物には以下の課題を学習させる。図2のような押しボタン()の付いたタスクパネルの両側にそれ

ぞれの動物を座らせ、双方が手前中央のボタンを押した時点で課題が開始する。その後、パネル中央の指示LED()が点灯した側のサルが左右どちらかのボタンを押し、それぞれのボタンに

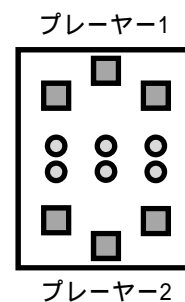


図2 課題装置

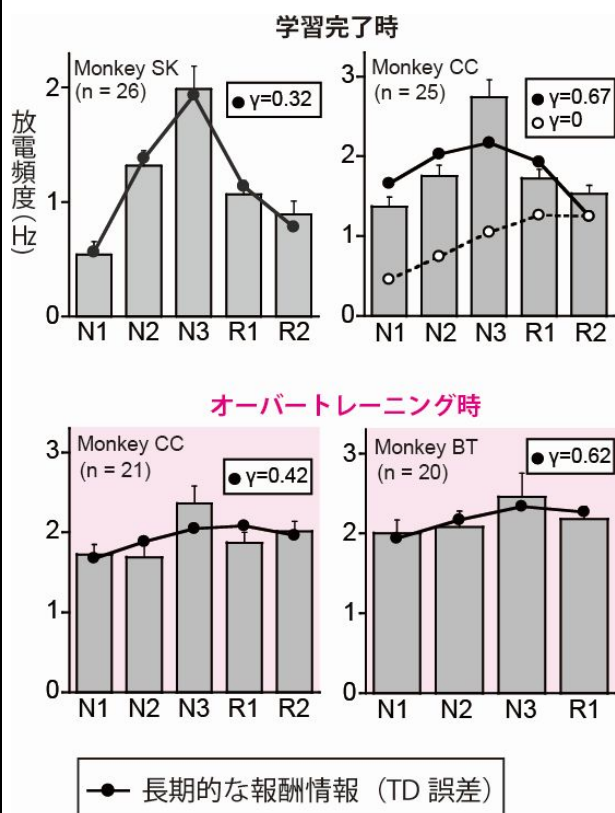
割当てられた確率に従って報酬(ジュース)が口元のパイプから与えられる。次の試行では反対側のサルが同様の選択行動を行い、それを交互に繰り返す。左右のボタンに割り当てられた報酬確率は課題1ブロック内(30~50試行)で一定であり([左-右]=[80%-20%][50%-50%][80%-20%])。プレイヤー間で共通しているため、動物は自分自身の行動結果に加えて、相手の行動とその結果からも報酬確率を推定することができ、相手の行動観察によって効率よく学習を進めることができる。また、プレイヤー間で報酬確率が独立している条件(相手の行動が参考にならない条件)と、一方のプレイヤーだけ、ある程度ブロックの確率配分を学習させた後に課題を行わせる条件(相手の行動がより信頼できる条件)を通常課題の間に組み入れ(条件の違いは指示LEDの点灯パターンで教示する)学習における相手の信頼性を変化させる。以上の行動課題を用いて、各試行における動物の選択行動とその結果から学習モデルを推定する。試行錯誤による動物の学習行動を記述するのに適していると考えられる強化学習理論やベイズ理論を応用して、自

分自身の行動とその結果のみを用いるモデル、他者の行動も利用するモデル、他者の行動とその結果両方を利用するモデルを比較し、どの学習モデルが最も良く実際の行動を推定できるのかを、課題条件ごとに調べる。新しいブロックの開始直後ほど他者の観察学習の頻度は高く、また、他者の行動の信頼性が高いほど、より他者の行動と結果を利用した学習を行っていることが予想される。学習モデルの推定には、動物の一試行ごとの行動パターンを用い、学習率や将来報酬割引率などのパラメータを計算する。その他にも、行動の反応速度やエラー頻度を計測し、学習へのモチベーションの指標とする。また、視点の位置や眼球運動速度を測定し、各プレーヤーがどこをどの程度見ているのかを正確に判別し、推定した学習モデルの裏付けとする（他者の信頼性が高いほど、相手の顔や手元をよく見るはずである）。更に、報酬到達直前のパイプの舌なめ運動から、動物の主観的な報酬期待の度合いを推定する。高確率で報酬が得られることが期待される場合には、強く、長い舌なめ運動が計測されると予想される。

4. 研究成果

(1) 3頭のニホンザルから、226個のドーパミン細胞の活動を記録し、課題遂行時の応答を調べた。ドーパミン細胞は各試行開始時の視覚刺激（条件刺激）に対して一過性の興奮応答を示した。課題学習完了時の放電頻度の大きさは、N1からN3にかけて報酬確率に伴って増大したが、ほぼ確実に報酬が得られるR1・R2では、むしろN3よりも減少していた。このことは、N1～N3開始時には、R1・R2の報酬も含んだ3回分の報酬価値を、R1開始時にはR2の報酬も含んだ2回分、R2開始時にはその試行1回分だけの報酬価値を（時間的に遠い報酬ほど価値を割り引いて）それぞれ予測しており、そのような複数の報酬情報が

細胞応答に反映されていることを示唆している。これらの応答をよく予測できるような将来報酬の割引率（ γ ）を計算機シミュレーションによって求めたところ、その値は動物の行動データ（行動結果提示前のパイプなめ行動）から算出した値とよく一致していた（図3上段）。また、予測誤差を表現すると考えられている、行動選択の結果（正解/不正解）を知らせる音刺激（強化因子）に対する応答は、同じ割引率の値を用いて計算した、



複数回の報酬情報を含んだ Temporal Difference (TD) 誤差の値との間に有意な相関が認められた。

図3 学習によるドーパミン細胞活動の変化

ここから更に学習を進めると、ドーパミン細胞応答は試行間の差が小さくなった（図3下段）。この応答は、次ブロック以降の報酬も含めた、4回以上の報酬情報を反映する強化学習モデルを用いて計算した TD 誤差によって良く説明できた。よって、オーバートレー

ニング時には、学習完了時より更に長い、ブロックをまたいだ超長期的な報酬情報を反映したドーパミン細胞活動が見られることが明らかになった。更に、中脳腹内側部のドーパミン細胞は背外側部の細胞よりも大きな放電頻度の応答を示し、かつ、より先の(の大きい)将来報酬情報を反映していることが確かめられ、ドーパミン細胞および線条体細胞活動の部位特異性を示唆する結果を得た。

(2) 3 頭のニホンザルを用いて、相互依存的な行動課題を学習させ、社会的状況において行動がどのように他者の行動や報酬獲得に影響を受けるのかを調べた。課題遂行中のサルの眼球運動や報酬が出る口元のパイプへの舌なめ運動を記録したところ、相手の行動依存的に報酬が得られる状況では、自分が選択行動をしていない時においても、相手が選択行動を行っている手の位置や、報酬をなめている口元などを注視し、報酬獲得を期待した舌なめ運動をしていることが確認でき、相手の行動とその結果をよく観察し、自身の行動に生かす戦略を採っていることが示唆された。

以上の結果から、社会的状況において動物が用いている戦略(学習モデル)や神経活動を明らかにしたことによって、社会のなかで適切な意思決定を行い、効率的に利益を獲得するための学習に関わる脳神経メカニズムの解明に貢献することが期待される。

5 . 主な発表論文等

〔学会発表〕(計1件)

(1) Enomoto K, Matsumoto N, Haruno M, Kimura M. サル中脳ドーパミン細胞によるオーバートレーニング時の部位特異的な将来報酬表現 Heterogeneous reward signals of midbrain dopamine neurons in over-trained

monkeys. 第37回日本神経科学大会.

2014.9.11. パシフィコ横浜, 神奈川県横浜市

〔図書〕(計1件)

(1) 山田洋, 榎本一紀. 中外医学社. Clinical Neuroscience Vol.32 No.1. 2014. 47-50

〔その他〕

ホームページ等

(1)<https://researchmap.jp/eno-kazuki/>
(2)<http://loop.frontiersin.org/people/181763/>

6 . 研究組織

(1)研究代表者

榎本 一紀 (ENOMOTO, Kazuki)

玉川大学・脳科学研究所・科研費研究員

研究者番号: 10585904

(2)研究分担者

なし

(3)連携研究者

なし