

平成 30 年 6 月 13 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K13505

研究課題名(和文)モデルフリー・モデルベースシステム間バランスを評価する小動物用行動試験の開発

研究課題名(英文) A novel behavioral task for investigating a mixture of model-free and model-based reinforcement learning in rodents

研究代表者

溝口 博之 (Mizoguchi, Hiroyuki)

名古屋大学・環境医学研究所・講師

研究者番号：70402568

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：適切かつ柔軟な意思決定を行うには、習慣的システムと目的指向的システムが状況に応じて切り替わりながら、協調的に働くことが重要である。この研究領域はヒトを対象とした研究が盛んであるのに対し、基礎研究(げっ歯類)では、システムバランスを評価する行動試験が未開発である。そこで本研究では、計算理論研究を取り入れ、げっ歯類を対象とした情報処理システムを評価する新規行動試験の開発を試みた。その結果、ラットは報酬量や報酬確率に依存した目的指向な行動選択を示したことから、タッチスクリーン弁別実験を立上げること成功した。しかし、目的とするモデルフィット化まで至らなかった。今後さらなる検討を行う予定である。

研究成果の概要(英文)：Decision-making and reinforcement learning arise from the involvement of two distinct processes that control choice, known as model-based and model-free strategies. Model-based choices are prospective, whereas model-free choices are retrospective. In general, normal performance depends on the flexible integration of model-based and model-free control. In this research, we employed a two-step sequential visual discrimination learning of complex stimuli in touch-screen task to establish a novel behavioral system for investigating a mixture of model-free and model-based reinforcement learning in rodent. Rats chose option on the reward amount- and provability-dependent manner. However, we failed to find best fitting behavioral modeling and parameters. We need to continue our experiments in future studies.

研究分野：神経薬理学

キーワード：モデルフリー・モデルベース 意思決定 行動選択

### 1. 研究開始当初の背景

脳には、意思決定や行動選択に必要な、習慣的システム（モデルフリー強化学習）と目的指向的システム（モデルベース強化学習）の2つの情報処理システムが存在する。習慣的システムは、直感的、連想的で、労力が掛からないことが特徴で、目的指向的システムは分析的、統制的で、労力が掛かることが特徴である。よって、適切かつ柔軟な意思決定には、両システムが状況に応じて切り替わりながら、協調的に働くことが重要である。一方、強迫的な過食症や薬物依存症患者は、暴飲暴食や薬物摂取という行為が自分にとって負の結果（肥満や依存）をもたらすと知りながら、その行動を選択する（Mizoguchi et al., PNAS, 2015）。コカインや覚せい剤投与は習慣的行動を増加させる事から、習慣的システムの過度な駆動が依存症などの連想的な反復行動を誘導すると考えられる（Zapata et al., J Neurosci, 2010）。そのような中、計算科学を組み込んだ意思決定試験において、習慣的システムと目的指向的システムを別個に抽出するのではなく、一緒に抽出することで、システム間の協調性やバランスを評価できる新しい行動タスクが開発された（Daw et al., Neuron, 2011）。このタスクを用いると、強迫障害患者や依存症患者は習慣的システム優位に選択行動するというシステム間のバランス異常が生じていることが示された（Voon et al., Mol Psychiatry, 2015）。しかし、責任脳領域や神経回路は不明であり、その脳領域や神経回路の破綻が精神疾患患者で見られる意思決定異常やシステム間のバランス異常を誘発するかどうかは分かっていない。

### 2. 研究の目的

本研究では、基礎研究において、システム間のバランスを評価できる強化学習タスクの確立に挑戦する。

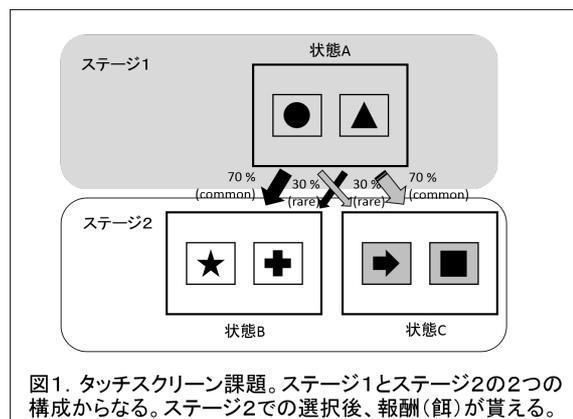
臨床試験で使用されたモデルフリーモデ

ルベース評価行動試験法を参考に、小動物用行動タスクの開発、強化学習モデルに基づくパラメータ推定を行うこととした。では、タッチパネルを用いた高度な視覚依存的餌弁別学習課題をプログラムする。このプログラムを用いて、実際にラットがヒトと同様に、モデルフリーとモデルベースの両方の情報処理システムを使って、行動選択するかどうか検討した。また、ラットの画像認知に対する基礎データを収集することにした。では、モデルフィット化により、weighting parameter(システム間比重) 固執、逆温度、学習係数などのパラメータ推定を行った。

### 3. 研究の方法

臨床試験で使用された評価タスクを参考（Daw et al., Neuron, 2011; Voon et al., Mol Psychiatry, 2015）に、タッチスクリーン選択課題とモデルフィット化を行った。

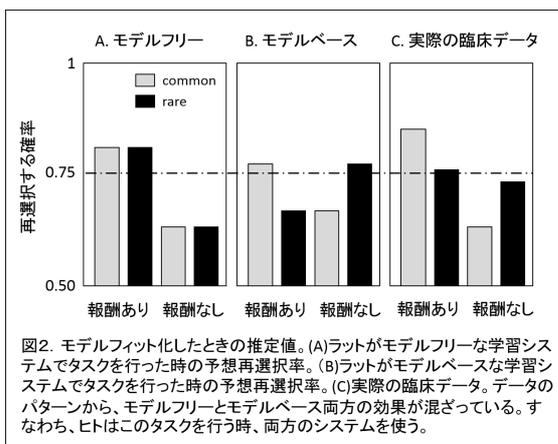
課題はステージ1とステージ2の2つのステージと3つの状態からなり、選択により状態の遷移を計算することで、モデルフリーとモデルベースを同時に抽出することが可能となる（図1）。



ラットはステージ1(状態A)でいずれかの図形を選択する。選択された図形に応じた確率で、ステージ2では状態Bもしくは状態Cのいずれかに遷移する。例えば、ステージ1で左の図形( )を選ぶと70%の確率で状態Bに、30%の確率で状態Cに遷移する。右の図形( )を選んだ場合は、その逆とな

る。ステージ2で選んだ選択肢(★、+、→、)に応じた報酬確率で強化子をもらえたりもらえなかったりする。モデルフリーな(状態の遷移を考慮しない)学習をすると、報酬があたえられたときにその直前の行動を繰り返す頻度が上がる。例えば、ステージ1で図形 を選び、30%のレアな確率で状態Cへの状態遷移が起こった時に、報酬が得られたとすると、モデルフリーな学習では、報酬が与えられたのでステージ1で選んだ図形の価値が上がり、次にその図形を選ぶ確率が高まる。

一方、モデルベースな(状態の遷移を考慮する)学習では、そのような状況で、価値の高まった状態Cに行くにはステージ1では左の図形より右の図形( )を選ぶと確率が高い、という知識(モデル)を利用する。つまり、そのような状況ではステージ1では を選択する確率が高くなる。



これらの予測をグラフにしたのが図2となる(参考文献を参考にした予想図)。これらはステージ1で選択した図形を次も選択する確率を縦軸で表している。「common」は70%の遷移が起きた場合で、「rare」は30%の遷移が起きた時である。

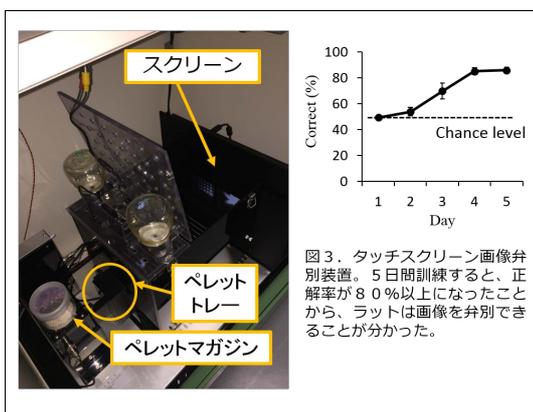
パネルAはモデルフリーな学習システムに依存した場合の例。ステージ1からステージ2への遷移がcommonだったかrareだったかに関係なく、報酬を得た時は再選択する確率が高まる。パネルBはモデルベースな学習システムに依存した場合の例。「common」な遷

移の後に報酬を得た場合は、stayする確率が高まるが、rareな遷移の後に報酬を得た場合は、stayする確率は低くなる(次は、そのステージ2の状態に高確率で遷移する、逆の選択をするようになる)。パネルCは実際の被験者の平均値を示している。モデルフリーとモデルベースの両方の効果が混ざったものになっていることから、ヒトは両方の計算システムを使うことを示している。

上記予測のもと、タッチスクリーン弁別装置を用いて、ラットに弁別学習を行わせた。

#### 4. 研究成果

タッチスクリーン行動解析装置を導入し、行動タスクの確立を目的に、タスクの作製とタスク難易度に対するラットのモチベーションなどを確認することにした。作製した課題では、ラットはステージ1(状態A)でいずれかの図形を選択する。選択された図形に応じた確率で、ステージ2では状態Bもしくは状態Cのいずれかに遷移する。先ず初めに、ステージ1と2で使用する2画像を1ペアとして、ラットが画像を弁別できるかどうか確認することにした。その結果、ラットの嗜好性が余り大きく影響されない3パターンのペア画像を見つけることができた(図3)。



次に、ステージ1からステージ2への遷移率およびステージ2における状態Bと状態Cの画像の報酬確率を100% vs 0%と50% vs 50%といったように、固定して行動試験を行った。その結果、ラットは試行を経験することで、確率が高い選択肢を選び、報酬を獲得するよ

うになることが分かった。

次に、タッチスクリーン行動解析システムの再現性と基礎データをとるため、ラットのタスク難易度に対するモチベーション、画像の遷移率、画像の逆転など課題に対する認知について詳細な検討を行うことにした。その結果、ラットの嗜好性が余り大きく影響されない3パターンのペア画像を見つけることができたことに加え、ラットは訓練を行うと、画像を認識し、報酬量が多い選択肢や報酬獲得確率が高い選択肢を選ぶことで、報酬を獲得できるようになることが分かった。特にラットの行動選択は、報酬獲得確率に強く依存することが分かった。また、予想していた以上に、ラットにとって、正解画像が逆転することが、課題遂行の難易度を上げることになることが分かった。

また、モデルフィット化したところ、状態Bと状態Cの報酬確率を固定するとモデルフリー、モデルベースが判別できないことが分かった。

以上のことから、タッチパネルを用いた高度な視覚依存的餌弁別学習課題をプログラムすることができた。また、ラットは報酬量や報酬獲得確率に依存して、画像選択を行うことが分かった。また、ステージ1からステージ2への遷移率およびステージ2における状態Bと状態Cの画像の報酬確率に依存して、報酬確率が高い選択肢を選べることが分かった。しかし、遷移率を固定するとモデルフリー、モデルベースが判別できないことが分かった。正解画像を逆転させると、ラットにとって課題遂行の難易度が上がることから、今後はこの事を考慮しつつ、例数を追加して遷移率を変動させ、ラットがヒトと同様に、モデルフリーとモデルベースの両方の情報処理システムを使って、行動選択するかどうかを検討する予定である。また、モデルフィット化について再検討を行う予定である。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計2件)

1. Takagishi Y, Katanosaka K, Mizoguchi H, Murata Y. Disrupted axon-glia interactions at the paranode in myelinated nerves cause axonal degeneration and neuronal cell death in the aged *Caspr* mutant mouse shambling. *Neurobiology of Aging*, 査読有, 43, 2016, 34-46.  
DOI:10.1016/j.neurobiolaging.2016.03.020.
2. Fujisawa H, Sugimura Y, Takagi H, Mizoguchi H, Takeuchi H, Izumida H, Nakashima K, Ochiai H, Takeuchi S, Kiyota A, Fukumoto K, Iwama S, Takagishi Y, Hayashi Y, Arima H, Komatsu Y, Murata Y, Oiso Y. Chronic hyponatremia causes neurologic and psychologic impairments. *Journal of the American Society of Nephrology*, 査読有, 27(3), 2016, 766-780.  
DOI:10.1681/ASN.2014121196.

[学会発表](計7件)

1. 溝口博之, 犬束 歩, 片平健太郎, 山田清文. リスク下の意思決定における島皮質 GABA 神経の役割. 日本アルコール・アディクション医学会学術総会, 2017. 9. 横浜.
2. 溝口博之, 山田清文. 依存性薬物がもたらす認知・意思決定異常. 第40回日本神経科学大会(日本神経精神薬理学会合同企画シンポジウム), シンポジウム, 2017. 7. 千葉.
3. Mizoguchi H, Yamada K. Nicotinic acetylcholine receptor as a target of harm reduction and treatment in addiction. 第51回日本アルコール・ア

ディクシオン医学会学術総会，シンポジウム，2016. 10. 東京.

4. 溝口博之，山田清文. 意思決定障害と創薬標的としてのニコチン受容体. 第 18 回応用薬理シンポジウム，シンポジウム，2016. 8. 名古屋.
5. 溝口博之，山田清文. 島皮質 GABA 神経は薬物依存ラットの意思決定に関与する. 第 39 回日本神経科学会，シンポジウム，2016. 7. 横浜.
6. 溝口博之，山田清文. 薬物依存症における意思決定異常. 第 46 回日本神経精神薬理学会，シンポジウム，2016. 7. ソウル.
7. 溝口博之. 薬物依存における意思決定と島皮質機能障害. 第 34 回日本生理心理学会大会，シンポジウム，2016. 5. 名古屋.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

ホームページ等

<http://www.riem.nagoya-u.ac.jp/center.html>

#### 6. 研究組織

##### (1) 研究代表者

溝口博之 (MIZOGUCHI, Hiroyuki)  
名古屋大学・環境医学研究所・講師  
研究者番号：70402568

##### (2) 研究分担者

該当なし ( )

研究者番号：

##### (3) 連携研究者

該当なし ( )

研究者番号：

(4) 研究協力者  
該当なし ( )