

## カテゴリー

---

脳科学, 運動学習

## タイトル

---

強化学習における遠心性コピーの役割 The role of efference copy in striatal learning

PubMed へ Fee MS et al : Curr Opin Neurobiol. 2014 Apr;25:194-200

## 内容

---

### 学習における基底核モデル

•強化学習とは、試行錯誤を通じて環境に適応する学習制御の枠組で、環境、行動、報酬を表す信号の収束が必要です。

•大脳基底核機能のモデルは、環境と報酬を表すシグナルの神経起源に関する十分に確立された仮説を持っているが、行動を表すシグナルの機能と起源はあまり明確ではない。

報酬について補足)

最近の研究でドーパミンを報酬とし、その報酬を予測をもとに目的指向的な行動を組み立てる役

割が分かってきた。ドーパミンは大脳基底核に強く作用する。最近、強化学習で学習信号に使わ

れる TD 誤差の振る舞いと類似した神経活動がサルの中脳 VTA や黒質網様部 (SNc) にあるドー

パミン作動 1 生ニューロンにおいて発見された。ドーパミン細胞の主な投射先は、大脳基底核の

線条体と大脳皮質前頭前野である。そこでも、報酬に関する予測的な反応が観測されている

•最近の知見は、探索的または可変的な行動が、中脳、脳幹および皮質における多様な「行動生成

回路]によって開始されることを示唆している。従って、学習するためには、線条体は、これらの行動生成回路で行われた作用決定のコピーを組み込む必要がある。

線条体について補足)

線条体は尾状核と被核の二つであらわされています。線条体は方法、何かのやりかたの記憶を司る場所で、「無意識かつ自動的かつ正確」「繰り返しの訓練で身に付く」という特徴がある。黒質からの運動に関わる神経伝達物質（ドーパミン）を受け、作動します。運動に関する情報の入力部とされています。

•遠心性コピー信号の役割を強調する強化学習の最近のいくつかの神経モデル、およびこれらの信号が入力信号と環境と報酬とにどのように統合されるかについてのアイデアを検討する。

•特定の環境で満足のいく効果を生み出す行動は、その状況になった時、再びその行動が起こる可能性がより高くなる。

補足) ソーンダイクの法則における効果の法則

• 満足をもたらす反応は結合を強くし、起こりやすくなる。(満足の法則)

• 不快をもたらす反応は結合を弱くし、起こりにくくなる。(不満足の法則)

• 満足や不快の程度が強いほど、結合力は大きくなる。(強度の法則)

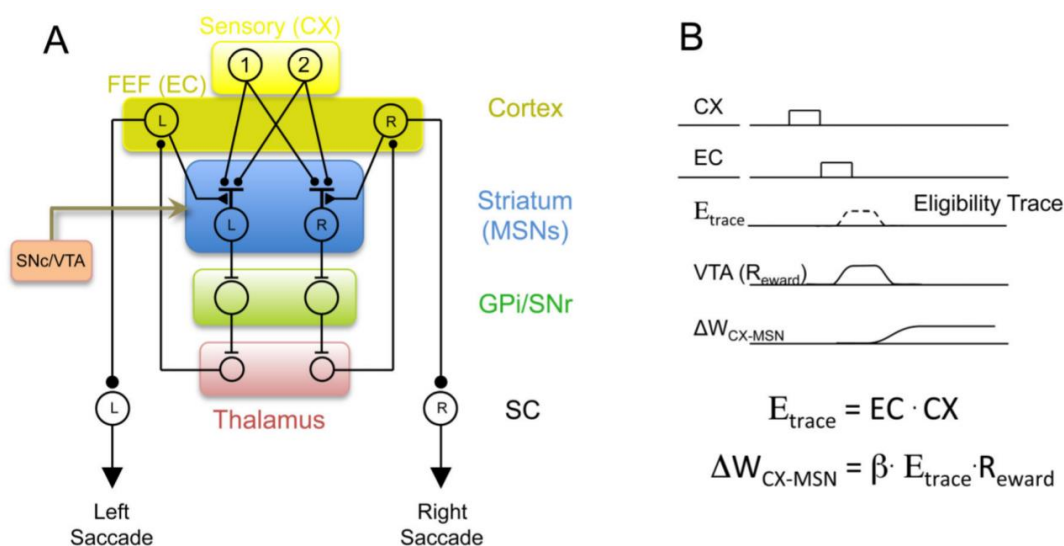
•最近の強化学習のいくつかの神経モデルは、遠心性コピー信号の役割を強調し、そのような信号を入力と統合して環境と報酬をシグナル伝達する方法についてのアイデアを取り入れている。

•基底核(BG)の神経回路は、学習行動の制御に関与することが知られており、BG の入力構造である線条体は、強化学習の神経実行の鍵となる構造として確立されている。

•この視点の最も魅力的なサポートの一部は、動物が訓練され、報酬を用いて、どの視覚刺激が提示されるかに応じて特定の方向にサッケードを作るための、眼球運動学習における基底核回路の役割を実証する作業から来る。

•これらの行動における BG 回路の役割のためのモデルの 1 つでは、報酬刺激の出現を表す皮質ニューロンは、尾状核の「直接経路」において中型有棘神経細胞 (MSN) を活性化すると考えられており、上丘のサッケード生成ニューロンを活性化して、報酬を与えられる方向に強いサッケードを引き起こす。

### 眼球運動学習モデル



(Fee MS et al : 2014)

PubMed へ

•運動行動の遠心性コピーを組み込んだ基底核機能のモデル(眼球運動学習モデル)。示されているのは、刺激とサッケード方向との間の関連の強化学習を実施するためのネットワークの概略図である。

•基底核機能のいくつかのモデルの一つでは、学習中に探索行動を生成する「自発的な活動」が線条体自体にあると仮定している。この場合、学習は簡単です。左または右にサッケードする決定がMSNの自発的な活動によって生成される場合、学習に重要な3つのシグナルはすべて、それぞれのシナプスで利用できます。環境信号はシナプス前入力によって皮質線条体のシナプスに運ばれ、行動の信号はシナプス後のスパイクによって運ばれ、報酬信号はドーパミンの相放出によって運ばれる。

•実際に、強化学習の根底にある皮質線条体の学習規則は、ゲートされた活動電位タイミング依存性シナプス可塑性(刺激頻度等に応じてシナプス可塑性の方向性が変化する現象)の一形態であることが示唆されている。この考えは、皮質線条体の可塑性におけるドーパミンの役割に関する最近の知見と一致している。

•基底核は学習後に特定の行動を引き起こすのに強力な役割を果たすようであるが、眼球運動学習の前または途中で探索行動の源ではない可能性があることが示唆されている。したがって、早期に学習する「探索的」サッケードは、線条体ではなく、上丘に投影され、サッケード生成を引き起こしたり影響を与えたりすることができる多くの脳回路の1つとして開始される可能性がある。

•より一般的には、未熟な訓練されていない動物によって生み出された自発的な行動は、運動野および脳幹に分布する無数の行動生成回路によって開始され得る。

•これらの回路は、外部の感覚刺激によって、または自発的な行動を促進する内因性の「ノイズの多い」メカニズムによっても活性化され得る。

例えば、若い鳥(ソングバード)の探索的なさえずりや鳴き声(歌声)の変化は、基底核回路を必要としない。むしろ、音声学習の根底にある変動性は、LMANとして知られる鳥類皮質の特殊回路内で駆動され、おそらくそれによって生成される。(AFP anterior forebrain pathway)

※ソングバードについて補足 1. 歌を学習するために、幼鳥は、まず父鳥が歌うのを聞き、ある程度の歌の記憶を脳内に形成します。少し経つと、幼鳥は自身で歌い始めます。練習している間、幼鳥は自身の歌を聞き、脳内にある記憶により良く符合するように、徐々に歌に磨きをかけます。このようにして、幼鳥は最終的に父鳥の歌を上手くコピーをすることができるのです。ソングバードのさえずりは、求愛やテリトリー宣言を意味し、個体ごとに個性がある。中脳ドーパミン系腹側被蓋野(VTA)内のニューロンが、雄が単独で歌っている時よりも、求愛のために歌う時のほうが、はるかに強く活性化されるという事実を発見しました。参照) 理研 BSI 鳥の歌(さえずり)の学習とコミュニケーションにおける脳のメカニズム 小鳥のさえずり学習の神経機構: 大脳基底核経路と強化学習モデル

## 学習についてのまとめ

•行動生成回路は、動物が取る可能性のある多くの行動のうちのどれを選択するか、相互阻害などの競合的メカニズムを組み込んでいる可能性が高い。代わりに、BG は、BG 以外の回路の動作を生成するためにすでに組み込まれている意思決定プロセスをバイアスする役割を果たしていることを示唆しています。その目的は、以前の経験が示した行動に有利にこれらの決定を傾けることです。

•最近の基底核機能のいくつかのモデルは、線条体の可塑性を形成するために、遠心性コピーのシグナルの潜在的な重要性を強調しています。

•競合的相互作用により、行動の1つが選択されるまで、異なる行動を表すニューロンの活動が蓄積され、行われたい行動を表すニューロンが抑制されます。選ばれたニューロンから線条体へのフィードバック信号は、どの線条体ニューロンが可塑性を受けられるかを決定する。学習後、基底核回路の出力は皮質回路にフィードバックされ、良好な結果として決定された方へ偏らせる。

•新しい行動は報酬ではなく、短期間の段階的ドーパミン作動性反応によって示される予期しない顕著な刺激の出現によって強化されることが示唆されています。これらの顕著な信号と環境および遠心性コピー信号との収束を用いて、特定の状況において新規な結果につながる行動を発見するモデルを提案した。

•最近では、ソングバードの歌声の強化学習の環境の中で、行動の遠心性コピーを利用した基底核機能のモデルが開発されている。皮質の可変生成の核のニューロン LMAN は歌の運動路に投射するだけでなく、歌関連の基底核回路領域 X に担保を送る。領域 X の MSN は、VTA（腹側被蓋野）/SNc からのドーパミン作動性入力によって領域 X に潜在的に伝達される、LMAN のバリエーション指令と歌声のパフォーマンスの尺度との間の相関を測定することが提案された。したがって、エリア X 内の MSN は、どの LMAN バリエーション指令がより良い歌声につながり、より悪い曲の歌声につながるかを発見することができる。

•もちろん、特定のバリエーションコマンドが発生する環境は重要です。

•LMAN によって駆動される歌声の器官の筋肉に対する付加的な緊張は、歌の中で一度に歌を良くするかもしれないが、異なる時にそれを悪化させるかもしれない。

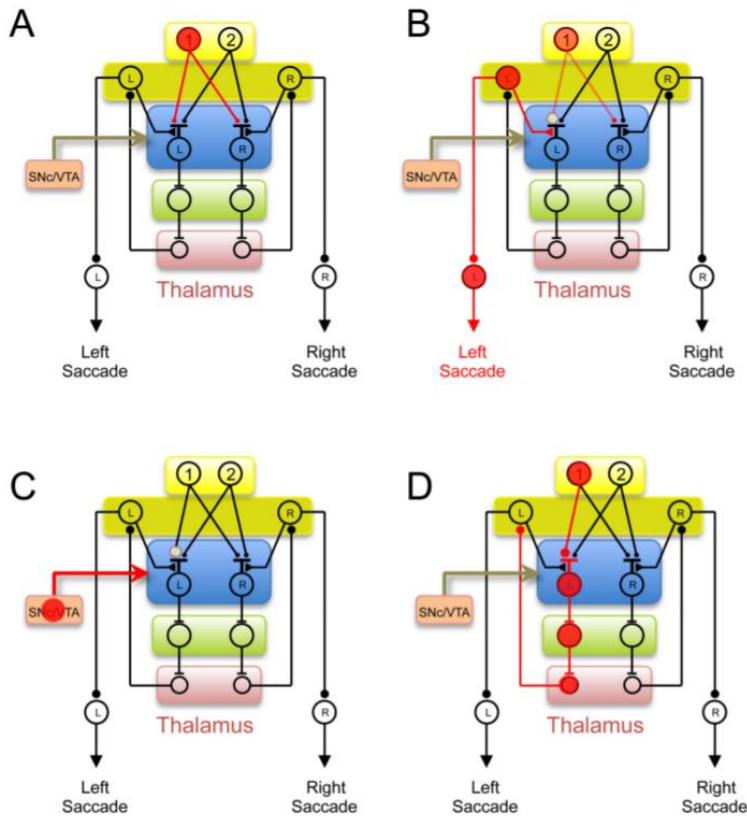
•このように、歌声のパフォーマンスの LMAN 活動との関係は、運動前野の核 HVC のニューロンから領域 X に送信される環境信号を利用することによって、独立して歌の各時間に実行されるべきである。これらのニューロンはまばらに活動的であり、それぞれが歌の少ない回数で急激な短いスパイクを生成し、集団として、歌全体にわたって活動的である可能性が高い。

•領域 X の各 MSN が、歌のすべての瞬間を表す HVC ニューロンの集合からシナプス入力を受け取り、LMAN からの遠心性コピー信号および歌声の評価信号と共に、各 MSN は強化学習を実施するために必要な環境、行動、報酬などが含まれすべての情報を有する。

- 先に説明したルールによれば、HVC 入力と LMAN の遠心性コピー入力の一致によって報酬（より良い歌）が続く場合、単純に HVC シナプスを強化すべきである。
- シナプスの適性な写しは、歌声の評価（報酬）信号の後の到着まで、HVC-LMAN の一致の「記憶」を維持することが想定される。
- 学習後、強化された HVC 入力は、LMAN ニューロンの活動が以前に観察されたその歌をより良く歌った時間に対応する時に MSN を強く駆動するであろう。
- この MSN の活動は、LMAN ニューロンが歌の際によりアクティブになるよう、順番に、視床への直接的経路を通して LMAN にフィードバックされなければならない。確かに、LMAN によって生成された変動性は、学習中にバイアスされ、誤った歌声を減少させる方向に押すという強い証拠がある。
- ソング学習のモデルに触発され、運動の遠心性コピーを使用するコンセプトは、最近、眼球運動学習の単純なモデルにまで拡張された。
- このモデルは、フランクと概念的に類似しているが、眼球運動線条体は、上丘の深層からのサッケード指令、または 前頭眼野(FEF)などの眼球運動の皮質領域から、組織化されて上丘に投影され、尾状核に側副投影を形成する。



•その考えは、猿が視覚的ターゲットと報酬につながるサッケードとの関連を学ぶ前に、FEF 回路がサッケードの変動性の源として機能し、おそらく、それぞれの視覚刺激に応答してランダムな「推測」を生成する、ソングバードの核 LMAN に類似しています。



上図の説明)

•学習ルールは次のようになります。

•シナプス前の環境入力の活性化は遠心性コピー入力後に報酬が続くとき、環境から MSN シナプスを強化する。

•例えば、刺激 1 の出現に続いて左にサッケードが現れた場合、この学習規則は、刺激 1 ニューロンから左サッケードチャンネルの MSN への接続のみを正確に強化する。

•学習後、この強化されたシナプスは、刺激 1 を表すニューロンが左 MSN におけるスパイクを駆動し、所望通りにサッケードの生成を左方向に偏らせることを可能にする。

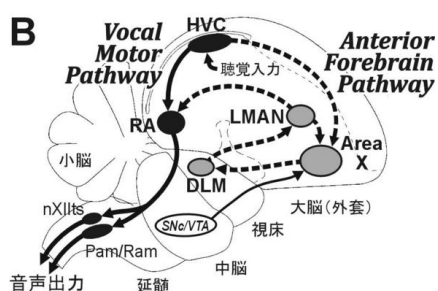
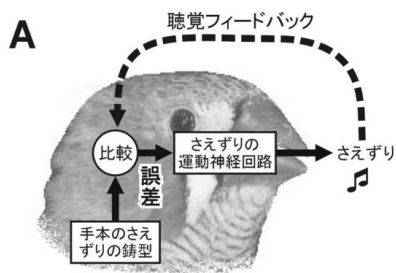
•この学習ルールの最終結果は、各環境 – MSN シナプスのシナプス強度が、強化学習の Q 学習モデルで想定されている状態 – アクション (Q) 値に似ていることである。

•報酬関連の可塑性は、MSN への環境入力でのみ生じ、一方、遠心性コピー入力は可塑性ではなく、可塑性をゲートする役割を果たす。MSN スパイクは、環境入力だけで駆動され、運動遠心性コピー入力では駆動されません。

## 私見・明日への臨床アイデア

•良くも悪くもリハビリは関われば何かしら学習を促すと思われるが、セラピストは良好な刺激は何かを考え、反応を見て、悪いものは抑制し、良好なものは良いことを共有し反復させ学習を強化させていくべきであると考えます。

•鳥の求愛ではないが、患者さんがセラピストを信頼し、リハビリを行うことは、学習にも大切になるのではと考えます。



氏名 Syuichi Kakusyo

職種 理学療法士