

# 運動意図の効率的な抽出を目的としたヒト運動制御機構の解明を目指した研究

研究代表者	池上 剛	情報通信研究機構	情報通信融合研究センター	研究員
共同研究者	ランディ フラナガン	クイーンズ大学	心理学部	教授
共同研究者	ダニエル ウォルパート	コロンビア大学	神経科学部	教授

## 1 はじめに

本研究は、2018年度長期海外研究援助、研究テーマ「運動意図の効率的な抽出を目的としたヒト運動制御機構の解明を目指した研究」の一環として行われた。

近年のAIブームに火をつけたDeep learningの中核を担うConvolutional Neural Networkは、脳の初期視覚野のニューロン特性や階層構造の知見をもとに開発された技術である。このように、極めて効率的な情報処理システムである脳に学ぶことは、ICT社会に革新をもたらす技術の開発につながる。

視覚系は脳の情報処理の中で最も理解が進んでおり、その知見は、ヒトの認知能力を遥かに凌ぐパターン認識能力をもつAIを可能にした。一方、運動系は未だ理解が不十分であり、その知見は、十分に活かされていない。現在のBrain Machine Interface (BMI)は、運動関連の脳活動を利用してロボットアームを操作することができるが、その動作の多様さや柔軟さは、ヒトの幼児にすら全く歯がたたない。

運動系の中でも特に、運動学習に関する脳内機序の深い理解は、脳活動から運動の意図を効率的に推定し、多様で柔軟な運動パターンを学習するBMIの開発につながる。このようなBMIは、四肢麻痺患者に、環境と自由に相互作用するためのツールを与えるだけでなく、全てのヒトの運動能力を増幅・拡張する、未来の人類にとっての（脳の情報通信を利用した）新しいICT技術になり得る。

本研究は、ヒトの運動学習を構成する代表的な2つの学習プロセスである、誤差学習と強化学習の関係を明らかにすることに焦点を置く。

## 2 強化学習と誤差学習の間の運動学習における共通性

### 2-1 背景

環境の変化に対して運動を適応させる能力は、我々ヒトの柔軟で巧みな行為にとって必要不可欠である。このような運動適応能力を支える代表的な2つの運動学習プロセスが誤差学習と強化学習である(1-5)。これら2つの学習プロセスは異なる計算機構と神経機構を持つと考えられている(6-8)。誤差学習は、大きさと方向のベクトル情報をもった誤差情報を利用して運動実行を更新するプロセスであり、主に小脳の機能と関連づけられてきた。一方、強化学習は成功・失敗といったバイナリ情報のようにスカラー表現された運動情報を利用して運動実行を更新するプロセスであり、主に大脳基底核の機能と関連づけられてきた。

運動学習の分野において、これらの学習プロセスは腕到達運動を用いて研究されてきた(9, 10)。被験者はロボットアームを操作することによってコンピュータ画面上のカーソルをターゲットに命中させる(図1)。その際、被験者の実際の運動とカーソルの運動の間にある新奇な視覚運動変換(11, 12)が加えられるため、学習初期は正確にターゲットに命中させることができない。しかし、課題の繰り返しによって視覚運動変換の学習が進み、ターゲットに命中させることができるようになる。

これまで多くの研究が、実際の手の位置に対して、画面に表示されるカーソルのゲイン(11, 13)あるいは方向(14-17)を変化させる視覚運動変換を用いて、運動学習機構を研究してきた。それらの研究のほとんどが誤差学習に焦点を当て、カーソルのフィードバックは連続的に或いは運動の終点にのみ与えられた。このカーソルのフィードバックによって、脳はベクトル表現された誤差情報を利用することができる。強化学習プロセスは、主に意思決定や認知コントロールといった認知研究の分野において盛んに研究されてきた(18-21)。運動分野における強化学習研究は最近注目を集めるようになってきた(1, 7, 22)。強化学習に焦点をあてた運動学習研究では、カーソルフィードバックは一切与えられず、課題が成功した場合にターゲットの色や音によって、課題の成否の情報を伝えるバイナリフィードバックのみが与えられた。

これまで誤差学習と強化学習の運動学習プロセスの関連を調べた研究は少ない。しかし、カーソルフィー

ドバックを終点にのみ与えた場合の誤差学習と強化学習の間に類似性があることを示唆する実験結果が存在する。Izawa と Shadmehr (1)は、運動方向の視覚運動変換課題を用いて、あるターゲットに対して課題を繰り返して獲得した学習効果が、未学習のターゲットにどの程度汎化するかを定量化した。すると、カーソルフィードバックを連続的に与えた誤差学習の場合、学習ターゲット以外の未学習のターゲットへも学習効果が広く汎化した。一方、強化学習の場合、未学習のターゲットへは学習効果がほとんど汎化しなかった。興味深いことに、カーソルフィードバックを終点にのみ与えた誤差学習の場合、連続的に与えた条件に比べて未学習のターゲットに対する汎化の範囲が狭まり、強化学習の結果に近くなった。この結果は、カーソルフィードバックを終点にのみ与えた場合の誤差学習と強化学習の学習プロセスが似ている可能性を示唆する。そこで、本研究は、連続的あるいは終点のみカーソルフィードバックを与えた誤差学習によって獲得した運動学習効果が強化学習へどのように転移するするかを調べた。もし終点のみカーソルフィードバックを与えた誤差学習と強化学習の間で共通した学習プロセスが存在するならば、学習効果が転移するはずである。この仮説を検証するために、我々は運動ゲインを変化させた視覚運動変換課題を用いて実験を行った。

## 2-2 方法

### (1) 被験者

右利き健常者 30 名 (24.2 ± 4.5 (mean ± s.d.) 歳 ; 女性 18 名、男性 12 名) を対象に実験を行った。本研究はコロンビア大学メディカルセンターの倫理委員会によって承認されており、ヘルシンキ宣言に準拠して行われた。全ての被験者は実験参加前に同意書に署名し、研究参加への同意を得た。被験者は無作為に 3 グループのうちの 1 つに割り当てられた (下記参照)。EL-C グループの 1 人の被験者は、予め設定された運動時間で運動を実施することができず、他の被験者に比べて約 3 倍もの実験時間がかかったため、以下の解析から除外した。

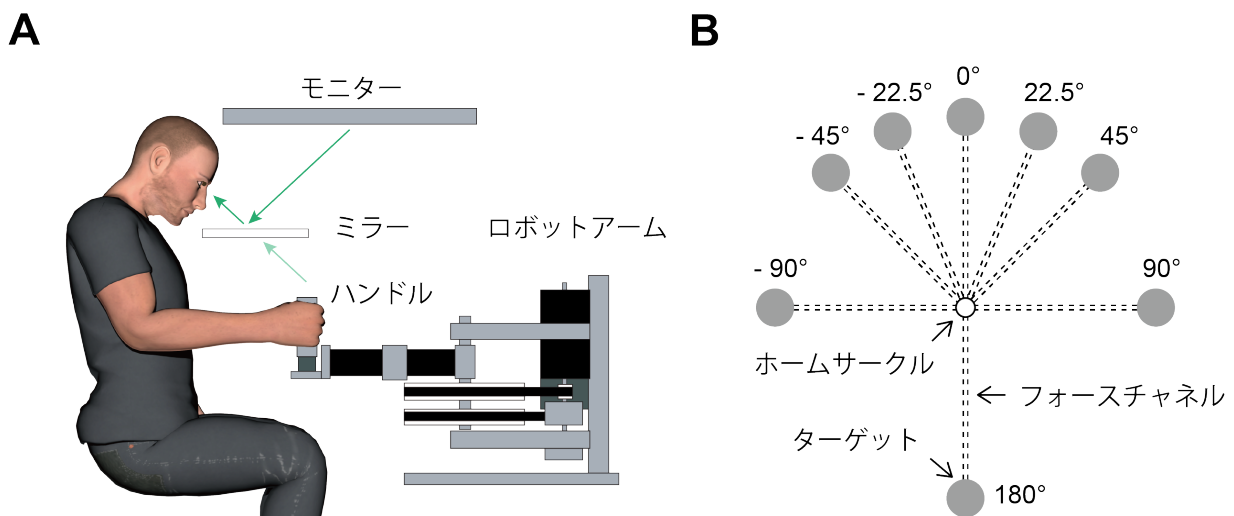


図 1 : 実験セットアップ (A) 被験者はロボットアーム (vBOT) のハンドルを右手で握り、到達運動を行った。ロボットアームの運動は平面上に拘束されており、トルクモータを介して力を生成することができる。上方に取り付けられたコンピュータモニターによって視覚刺激と運動のフィードバックが表示され、被験者はミラーを介してそれらを見ることができた。(B) 被験者は各試行において 8 つのターゲットのうちの 1 つに対して到達運動を行った。0°方向ターゲットはトレーニング試行 (新奇ゲインに適應するための試行) とプローブ試行 (汎化関数を評価するための試行) の両方に使用された。一方、他 7 つのターゲットはプローブ試行のみに使用された。全ての試行は、vBOT によって生成されたフォースチャネル環境下で行われた。

### (2) 実験装置

被験者は椅子に座り、右手でロボットアーム (vBOT) (23) を操作して、腕到達運動課題を行った。(図 1) 1000Hz で被験者の手先 (ハンドル) の位置と速度が計測された。右手前腕はエアースレッドによって支えられ、手と腕の動きは平面に拘束された。ターゲット (半径 1cm)、ホームサークル (半径 0.5cm)、カーソル (半

径 0.3cm) がコンピュータ画面上に提示された。これらの視覚刺激はミラーを介して手の運動平面上に投影された。また、ミラーによって被験者は自分の手やロボットアームを直接見ることはできなかった。

### (3) 実験手続き

被験者は胸の位置から 35 cm 前方に位置するホームサークルからカーソルを動かしてターゲットへの到達運動を行った。ターゲットは 8 つの位置のうちどこか 1 つの位置に提示された。ターゲットはホームサークルから 10 cm 離れており、 $0^\circ$  (正中矢状面前方)、 $\pm 22.5^\circ$ 、 $\pm 45^\circ$ 、 $\pm 90^\circ$ 、and  $180^\circ$  (時計回りの角度が正の値をとる) の位置に配置された。

視覚運動ゲインの変化は、ホームサークルから放射状方向に沿って、カーソル位置に対する実際の手の位置に対して付加された。例えば、ゲインが 1.33 の場合、10cm 離れたターゲットに到達するためには被験者は 7.5 cm ( $\sim 10/1.33$ ) の到達運動を行う必要がある。ゲインが 1 の場合、カーソルは実際の手の位置に表示される。全ての試行は、バネ様の力 (バネ係数:  $3000 \text{ N m}^{-1}$ ; 減衰係数:  $5 \text{ Ns m}^{-1}$ ) を利用したフォースチャンネル環境下(24)において行われ、手の動きがホームサークルからターゲットを結ぶ直線状に拘束された。これによって、ゲイン適応に対するあらゆる行動変化を、方向ではなく距離に対する運動制御能力に起因させることが可能になった(25)。

各試行はカーソルがホームサークルで静止した状態 ( $3 \text{ cm s}^{-1}$  で 0.1 秒間) から開始された。ターゲットは運動開始を告げる音と同時に提示された。もし、被験者の運動が 0.8 s 以上かかった場合は、“move fast” という警告メッセージが提示され、その試行が再度繰り返された。各試行終了後には、ロボットアームが自動的に被験者の手をホームサークルへ動かした。手がホームサークルへ戻される間は、カーソルフィードバックは与えられなかった。

実験中の試行は、トレーニング試行とプローブ試行の 2 種類に分類された。トレーニング試行では、被験者はパフォーマンスに関するフィードバックが得られ、ターゲット方向は常に  $0^\circ$  方向であった。プローブ試行は 8 つのターゲットに対する運動学習の汎化を測定するために用いられた。プローブ試行では、フィードバックは一切与えられなかった。

各試行に与えられるフィードバックは以下の 4 種類に分類された。

連続エラーフィードバック (Continuous error feedback): 運動中、カーソルは常に表示される。カーソル中心をターゲット内で停止することができた場合、チャイム音とともにターゲットが爆発するアニメーションが提示され、課題の成功がフィードバックされる。失敗時にはそれらのフィードバックは与えられない。

終点エラーフィードバック (Endpoint error feedback): 運動が開始して、手がホームサークルの外に出た瞬間にカーソルのフィードバックが消える。運動の終点において 0.2 秒間静止した後に、再びカーソルが提示される。課題の成否に関するフィードバックがターゲットの爆発アニメーションとチャイム音によって通知され、失敗時にはそれらのフィードバックは与えられない。

強化フィードバック (Reinforcement feedback): カーソルはホームサークル内でのみ表示され、運動中も運動後も表示されない。課題の成否に関するフィードバックがターゲットの爆発アニメーションとチャイム音によって通知され、失敗時にはそれらのフィードバックは与えられない。

フィードバックなし (No feedback): カーソルはホームサークル内でのみ表示され、運動中も運動後も表示されない。課題の成否に関するフィードバックも一切与えられない。

フィードバックなしの試行タイプはマゼンタ色のターゲットによって示され、それ以外の試行タイプは全て黄色のターゲットによって示された。

### 実験グループ

被験者は 3 つのグループ (各グループ 10 人) のうちのいずれかにランダムに割り当てられる。それらのグループは、徐々に変化していく視覚運動ゲインに適応するトレーニングフェーズ中に与えられるフィードバックタイプが異なる。EL-C (Error-based learning with continuous feedback)、EL-E (Error-based learning with endpoint feedback)、RL (Reinforcement learning) グループはそれぞれ、連続エラーフィードバック、終点エラーフィードバック、強化フィードバックが与えられた。

### 試行の構造

実験装置や実験内容に慣れることを目的として、被験者は本実験の前に全ての試行タイプを含んだ馴化セ

セッションを行った。馴化セッションは 356 試行から構成され、8 つのターゲットに対して 1 のゲインで試行を行った。被験者には、全ての試行タイプにおいて、類似したベースライン運動（1 のゲイン条件下での運動）が実行されるように、ピーク速度の範囲を 40 ~ 60 cm s<sup>-1</sup> に設定し、範囲外の場合は“too slow”あるいは“too fast”のメッセージが提示され、その試行は再度繰り返された。このピークスピードの制限は馴化セッションのみで設けられた。その後、練習セッションが行われた。練習セッションは本実験と同じ内容の短縮版として、ただし 1 のゲインで行われた（下記参照：本実験の説明後に記載）。

本実験では、被験者は最初にトレーニングフェーズ（training phase）を行った。被験者は 0° 方向ターゲットに対するトレーニング試行を行い、フィードバックタイプは被験者グループによって異なった。被験者は、1 ブロックあたり 15 回のトレーニング試行の試行ブロックを 21 ブロック行い、その際 1 ブロック毎にゲインが 0.167 増加し、最終的に 1.33 となった。最後に、1.33 のゲインでさらに 2 ブロックが行われた。つまり、トレーニングフェーズは全 23 ブロックで構成された。

その後、被験者は、ジェネラリゼーションフェーズ（generalization phase）を行った。1 ブロックあたり 24 試行の試行ブロックを 10 ブロック行った。各ブロックにおいて、被験者は 2 回のトレーニング試行の後に 1 回のプローブ試行が続く、3 試行一組の 3 連試行を 8 つのターゲットのそれぞれに行った（合計 24 試行）。ターゲットの順番は各ブロックにおいて疑似ランダム化された。

ジェネラリゼーションフェーズの後、被験者は 15 試行のトレーニング試行（0° 方向ターゲット）からなるポストジェネラリゼーションフェーズ（post-generalization phase）を行った。

最後に、被験者は、ポストジェネラリゼーションフェーズの後にトランスファーフェーズ（transfer phase）を行った。EL-C と EL-E の 2 つのグループは、強化フィードバックによる 60 回のトレーニング試行を 1.33 のゲインで行った。

練習セッションは、各被験者が本実験で経験する 2 種類のフィードバック条件下において、1 のゲインで試行を行った。被験者は、30 回のトレーニング試行の後、5 ブロックのジェネラリゼーションフェーズ（24 試行/ブロック）を行った。

#### （４）データ解析

本研究は、トレーニングフェーズ後のジェネラリゼーションフェーズと誤差フィードバックから強化フィードバックへの学習の転移に焦点を当てて解析する。

各試行において、ターゲット距離（10cm）の手の運動距離（ホームサークル中心から運動終了時の手の位置）に対する比率として運動ゲインが計算された。

ジェネラリゼーションフェーズのトレーニング試行中、RL グループの被験者の手の運動は、ターゲット（報酬ゾーン）外へとドリフトし、ターゲット範囲に戻れなくなることがあった。カーソルフィードバックが得られる EL-C と EL-E グループでさえ、トレーニング試行においてターゲット範囲内の運動を実することができない場合があった。新奇環境への適応した運動の汎化能力を評価することが目的であるため、我々は以下の手続きを採用し、運動適応が認められない試行を解析から除外した。ジェネラリゼーションフェーズで繰り返し行う 3 連試行（2 回のトレーニング試行と 1 回のプローブ試行）に対して、2 回目のトレーニング試行で課題が不成功であった場合、その 3 連試行内のプローブ試行を課題から除外した。この手続きによって、EL-C、EL-E、RL グループの 80 回のプローブ試行のうち 5.9%、20.4%、45.0%が除外された。ただし、このような試行の除外を行わなかったとしても、汎化を評価するために行われた ANOVA 解析の有意性の解析結果は同じであったこと付け加えておく。

汎化を評価するために、各ターゲットに対してのプローブ試行の運動ゲインの試行間平均が計算された。各ターゲットの運動ゲインがトレーニングターゲット（0° 方向）への運動適応に対する割合によって表現されるように標準化された。つまり、完全な汎化は 1 で表された。仮に、被験者がトレーニングターゲット（0° 方向）に対して完全に適応した（つまり、運動ゲイン=1.33）が、あるプローブターゲットに対してゲインの変化が全くなかった場合（運動ゲイン=1）、汎化指標は 0.75（ $\sim 1/1.33$ ）となる。

我々は、グループ間の汎化の差を調べるために、2 要因（グループ×ターゲット）混合計画分散分析を行った。球面性の仮定が成立しない場合は、グリーンハウス・ゲイザー法による補正を行った。有意水準は 0.05 に設定した。さらに、トレーニングターゲットから偏差角の関数として標準化されたゲインに対して以下のガウシアン関数によってフィッティングを行い、ゲイン変化を定量化した。

$$g(\theta) = a \cdot \exp \frac{\theta^2}{2\sigma} + (1-a) \quad (1)$$

$a$  と  $\sigma$  は汎化関数の振幅と幅を表し、 $\theta$  はトレーニングターゲットからの偏差角である。

### 2-3 結果

3 つの被験者グループが、0°方向（前方）ターゲットに対して到達運動を行っている間に、視覚運動ゲインの変化（1 から 1.33 まで）が徐々に付加された。このトレーニングフェーズの間、各グループは、連続的なカーソルフィードバック (EL-C)、運動終点のみのカーソルフィードバック (EL-E)、あるいはカーソルフィードバックはなく課題の成否情報のみの強化フィードバック (RL) が与えられた。

このゲイン変化は非常にゆっくり（15 試行毎に 0.167 だけ増加）であったため、被験者は視覚運動ゲインの変化に気づけなかった。しかし、全てのグループの被験者は、運動距離を徐々に短くすることによって、このゲイン変化に上手く適応することができた（図 2A-C のトレーニングフェーズ）。このような運動適応の潜在性（無意識的な運動の変化）は、様々な課題において認められる運動学習の特徴の一つである（26, 27）。

その後のジェネラリゼーションフェーズにおいて 0°方向のトレーニングターゲットに対して適応した運動学習効果が、他のターゲットに対してどの程度汎化するのかを調べた（注：図 2A-C はトレーニングターゲットに対する試行データのみが描画されている）。図 1D は 3 つのグループの汎化パターンを示している。2 要因 ANOVA の結果、グループ ( $P = 0.028$ ) とターゲット方向 ( $P = 0.001$ ) の要因において有意な主効果が認められたが、有意な相互作用はなかった ( $P = 0.679$ )。汎化パターン（被験者平均）は、ガウシアン汎化関数の標準偏差 (SD: 式 1 の  $\sigma$ ) と振幅 (式 1 の  $a$ ) を持つ 2 パラメータモデルでフィッティングされた。SD はグループ間で類似していた (EL-C: 18.33; EL-E: 18.26; RL: 21.36) 一方、振幅は大きな違いがあった。RL

(0.08) の振幅は EL-C (0.18) よりも小さく、RL の方が EL-C に比べてより広い汎化パターンを持つことが示唆された。また、EL-E (0.12) は RL と EL-C の中間的な値を示した。この結果は、終点フィードバックによる誤差学習は、連続的フィードバックによる誤差学習だけでなく強化学習とも学習プロセスを共有している可能性を示唆する。

強化学習の運動学習メカニズムを調べた先行研究（28, 29）の結果を考慮すると、RL グループで観察されたより広い汎化パターンは、被験者が認知的戦略（つまりターゲットの手前を狙うなど）（15, 30）を利用することによって、未学習のターゲットへの部分的な汎化を実現していた可能性を反映しているかもしれない。もしそうだとすれば、終点フィードバックもまた、そのような認知的戦略を促進したのかもしれない（31）。

ジェネラリゼーションフェーズ直後には、0°方向ターゲットへの 15 回のトレーニング試行から構成される短いポストジェネラリゼーションフェーズが行われた。そして最後に、2 つの EL グループに対するトランスファーフェーズの振る舞いが調べられ、1.33 のゲインのまま、フィードバックが強化学習フィードバックに変更された。もし、終点フィードバックによる誤差学習が強化学習と学習プロセスを共有しているとすれば、EL-C グループに比べて EL-E グループのほうが、より学習が転移するはずだ。

図 2A, B に示されるように、EL-C グループの被験者は強化学習フィードバックに切り替わると、ゲインへの適応を保つことができなくなると一方、EL-E グループの被験者は適応を保つことができた。この効果を定量化するために、各被験者に対して、ポストジェネラリゼーションフェーズの最後 15 試行からトランスファーフェーズの 2 ブロック目（フィードバックの切り替えによるドリフト効果が安定した 16 試行から 30 試行目）までの手の運動距離の変化が計算された。すると、この変化は EL-E グループよりも EL-C グループの方が大きかった (t-test,  $P = 0.022$ )。さらに、我々は 2 つのグループのトランスファーフェーズの 2 ブロック目を比較したところ、EL-E グループ（9 人中 6 人が報酬ゾーンの外）に比べて EL-C グループ（10 人中 2 人のみが報酬ゾーンの外）の方が手の運動距離は大きな値を示した (t-test,  $P = 0.011$ )。

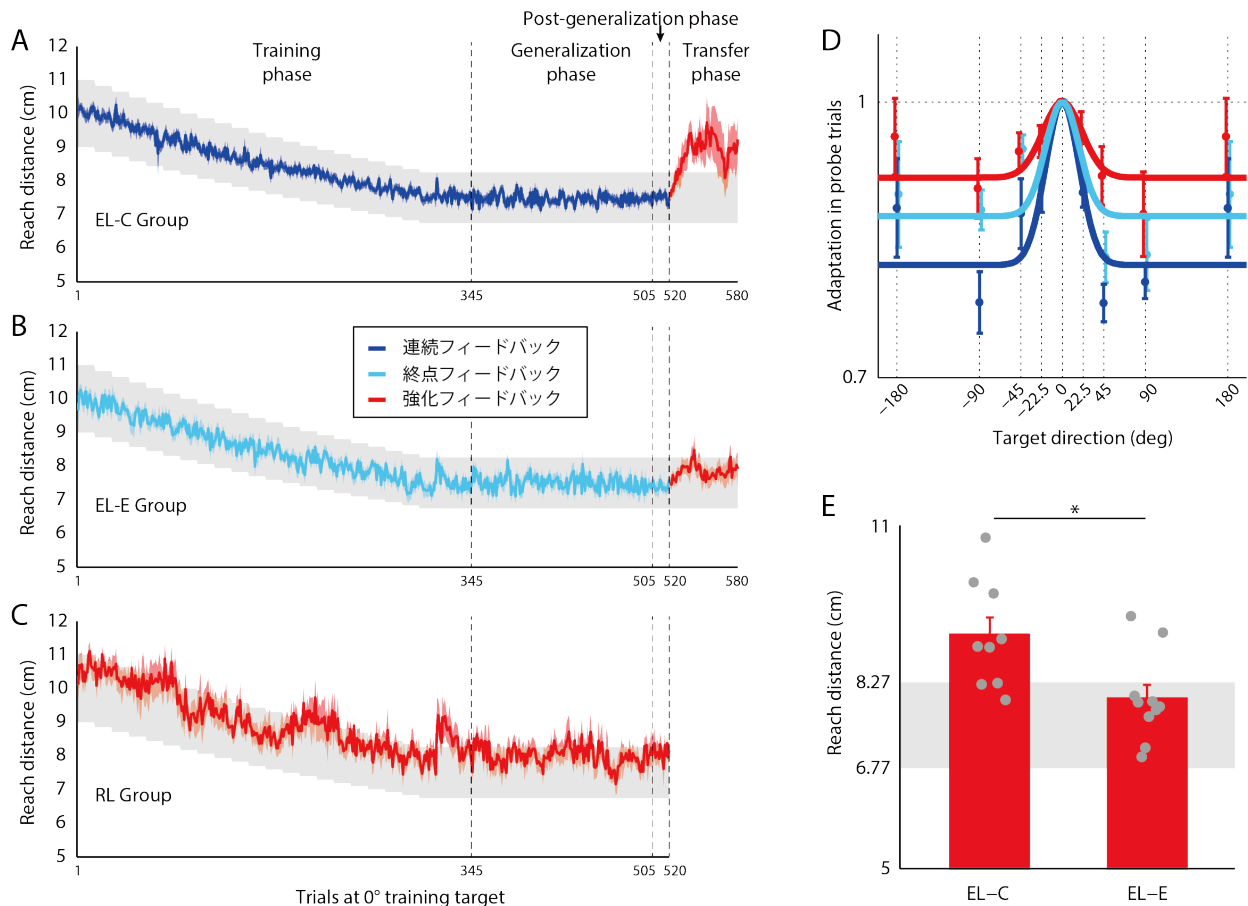


図 2：実験結果 (A-C) EL-C (青色), EL-E (水色), RL (赤色) グループの 0°方向ターゲットへの試行に対する手の運動距離 (reach distance)。 (D) プローブ試行 (フィードバックなし) におけるターゲット方向の関数としての運動適応の汎化。最初に、各プローブターゲットに対する運動適応を各被験者内の試行間平均として産出された。次に、その各プローブターゲットの運動適応は、トレーニングターゲット (0°方向) へのプローブ試行の運動適応によって標準化された。図中には標準化された運動適応の被験者間平均が描かれている。色付きのサークルは被験者平均を、エラーバーは標準偏差を示す。 (E) EL-C と EL-E グループのフィードバックが強化フィードバックに切り替えられた後のトランスファーフェーズの 2 ブロック目 (16~30 試行目) における手の運動距離。灰色のゾーンは手の位置に対するターゲット範囲 (報酬ゾーン)。手の位置に対するターゲット範囲はゲイン変換の影響を受けて変化 (1.33 のゲインに対して直径約 1.5 cm) するが、カーソルの位置に対するターゲット範囲は一定 (直径 2 cm) である。

## 2-4 議論

2つの代表的な学習機構である誤差学習と強化学習は、一般的に異なる計算機構と異なる神経基盤をもつ、異なるプロセスだと考えられてきた。運動学習分野においては、これらの2つの学習機構は到達運動課題を用いて調べられてきた。特に、ターゲットにカーソルを命中させるために、新奇環境に対してどのように手の運動が適応・修正されるかが調べられてきた。我々は、視覚運動ゲイン課題を用いて、本研究は運動の終点でカーソルフィードバックが与えられる誤差学習と、運動の終点で課題の成否に関するフィードバックが与えられる強化学習の間で、学習プロセスが共有されているという仮説を検証した。

被験者があるターゲットに対して繰り返し運動を行い、徐々に変化する視覚運動ゲインに対して十分に適応した後、その運動適応の未学習のターゲットに対する汎化パターンを調べた。すると、強化学習は学習ターゲットから離れた未学習のターゲットに対してより広く大局的な汎化パターンを示した一方、連続的なカーソルフィードバックを与えた誤差学習は学習ターゲットを中心により局所的な汎化パターンを示した。興

味深いことに、終点のみにカーソルフィードバックを与えた誤差学習の汎化パターンは、両者の中間のパターンを示した。本研究にとって最も重要な解析として、2つの誤差学習グループ（EL-CとEL-E）に対して、誤差フィードバックから強化学習フィードバックに切り替わったときに、学習が転移するかどうかを調べた。すると、終点フィードバックから強化学習フィードバックには学習の転移が生じた一方、連続フィードバックに対しては学習の転移は生じなかった。

これら結果は、終点フィードバックによる誤差学習と強化学習には共通の学習プロセスが関与することを示唆する。連続的フィードバックによる誤差学習は一般的に内部モデル(32)の学習と関連づけられる。内部モデルの学習を通じて、変化した手の視覚フィードバックに適合するように、運動指令と固有感覚知覚の両方が変化する。しかし、終点フィードバックによる誤差学習はこの内部モデルの学習にあまり依存していないかもしれない。その代わりに、強化学習と同様に、新しい行為と結果の関係 (action-outcome relationship) の学習により依存しているかもしれない(2, 28, 29)。運動中に視覚フィードバックを得ることができない終点フィードバックによる誤差学習と強化学習は、到達運動を制御するために固有感覚を利用したフィードバックコントロールポリシーを共に使っているかもしれない(31)。対して、連続的フィードバックによる誤差学習は、視覚フィードバックにより依存したコントロールポリシーを使っているのかもしれない。そのために、強化フィードバックが与えられると、その学習効果が転移しなかったと考えられる。

終点フィードバックによる誤差学習の汎化パターンが、連続フィードバックによる誤差学習と強化学習の中間値を示すという結果は、回転の視覚運動変換課題を用いた Izawa と Shadmehr(1)の結果と類似していた。しかし、Izawa と Shadmehr らの結果では、強化学習がより局所的な汎化パターン、連続フィードバックによる誤差学習がより大局的な汎化パターンを示した。これらの結果の違いは、回転変換とゲイン変換という課題の違いによる可能性がある。もしくは、Izawa と Shadmehr の研究では、強化学習条件において、被験者が視覚運動変換に十分に適応できていなかった可能性もある。彼らは、10 cm離れたターゲットに対する到達運動に対して、最大 8° の回転変換を付加した。その回転変換に適応するには、被験者は運動の終点を約 1.4cm 修正する必要がある。一方、我々は同じ 10 cm離れたターゲットに対する到達運動に対して、最大 0.33 のゲイン変換を付加した。そのゲイン変換に適応するには、被験者は運動の終点を約 2.5cm 修正する必要がある。つまり、本研究の方が、彼らの先行研究に比べて、約 80%も学習を多く誘起している。さらに、Izawa と Shadmehr の研究では、強化学習が誤差学習に比べて運動のばらつきが有意に大きく、強化学習の適応後の運動が適応前の運動と有意に差があるかどうかは不明である。一方、我々の研究では、RL グループの適応後の運動距離は適応前に比べて有意に変化していた ( $p < 0.05$ )。いずれにせよ、強化学習と誤差学習の汎化パターンは、今後の研究によって詳しく調べる必要がある。

最後に、これまで誤差学習と強化学習は全く異なる学習形式だと考えられてきた。しかし、本研究結果は誤差フィードバックが終点のみに与えられる場合、両運動学習プロセスは、学習プロセスを部分的に共有していることを示した。

## 【参考文献】

1. Izawa J, Shadmehr R. Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol.* 2011;7(3):e1002012.
2. Therrien AS, Wolpert DM, Bastian AJ. Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain.* 2016;139(Pt 1):101-14.
3. Therrien AS, Wolpert DM, Bastian AJ. Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. *eNeuro.* 2018;5(3).
4. Cashaback JGA, Lao CK, Palidis DJ, Coltman SK, McGregor HR, Gribble PL. The gradient of the reinforcement landscape influences sensorimotor learning. *PLoS Comput Biol.* 2019;15(3):e1006839.
5. Cashaback JGA, McGregor HR, Mohatarem A, Gribble PL. Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Comput Biol.* 2017;13(7):e1005623.
6. Doya K. Reinforcement learning in continuous time and space. *Neural Comput.* 2000;12(1):219-45.

7. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor Learning. *Compr Physiol.* 2019;9(2):613-63.
8. Wolpert DM, Diedrichsen J, Flanagan JR. Principles of sensorimotor learning. *Nature Reviews Neuroscience.* 2011;12(12):739-51.
9. Shadmehr R, Wise SP. *The computational neurobiology of reaching and pointing.* Cambridge, Massachusetts: The MIT Press; 2005.
10. Shadmehr R, Smith MA, Krakauer JW. Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci.* 2010;33:89-108.
11. Krakauer JW, Pine ZM, Ghilardi MF, Ghez C. Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J Neurosci.* 2000;20(23):8916-24.
12. McDougle SD, Ivry RB, Taylor JA. Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends Cogn Sci.* 2016;20(7):535-44.
13. Pearson TS, Krakauer JW, Mazzoni P. Learning Not to Generalize: Modular Adaptation of Visuomotor Gain. *Journal of Neurophysiology.* 2010;103(6):2938-52.
14. Taylor JA, Ivry RB. Flexible cognitive strategies during motor learning. *PLoS Comput Biol.* 2011;7(3):e1001096.
15. Taylor JA, Krakauer JW, Ivry RB. Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *The Journal of Neuroscience.* 2014;34(8):3023-32.
16. Ikegami T, Hirashima M, Taga G, Nozaki D. Asymmetric transfer of visuomotor learning between discrete and rhythmic movements. *J Neurosci.* 2010;30(12):4515-21.
17. Ikegami T, Hirashima M, Osu R, Nozaki D. Intermittent visual feedback can boost motor learning of rhythmic movements: evidence for error feedback beyond cycles. *J Neurosci.* 2012;32(2):653-7.
18. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997;275(5306):1593-9.
19. Schultz J, Friston KJ, O'Doherty J, Wolpert DM, Frith CD. Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron.* 2005;45(4):625-35.
20. Ribas-Fernandes JJ, Solway A, Diuk C, McGuire JT, Barto AG, Niv Y, et al. A neural signature of hierarchical reinforcement learning. *Neuron.* 2011;71(2):370-9.
21. Botvinick MM, Niv Y, Barto AC. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition.* 2009;113(3):262-80.
22. Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW. Overcoming motor "forgetting" through reinforcement of learned actions. *J Neurosci.* 2012;32(42):14617-21.
23. Howard IS, Ingram JN, Wolpert DM. A modular planar robotic manipulandum with end-point torque control. *J Neurosci Methods.* 2009;181(2):199-211.
24. Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA. Persistence of motor adaptation during constrained, multi-joint, arm movements. *J Neurophysiol.* 2000;84(2):853-62.
25. Pine ZM, Krakauer JW, Gordon J, Ghez C. Learning of scaling factors and reference axes for reaching movements. *Neuroreport.* 1996;7(14):2357-61.
26. Schmidt RA, Lee TD. *Motor control and learning : a behavioral emphasis.* 3rd ed. Champaign, IL: Human Kinetics; 1999. xvi, 495 p. p.
27. Orban de Xivry JJ, Criscimagna-Hemminger SE, Shadmehr R. Contributions of the motor cortex to adaptive control of reaching depend on the perturbation schedule. *Cereb Cortex.* 2011;21(7):1475-84.
28. Codol O, Holland PJ, Galea JM. The relationship between reinforcement and explicit control during visuomotor adaptation. *Scientific reports.* 2018;8(1):9121.
29. Holland P, Codol O, Galea JM. Contribution of explicit processes to reinforcement-based motor learning. *J Neurophysiol.* 2018;119(6):2241-55.
30. McDougle SD, Bond KM, Taylor JA. Explicit and Implicit Processes Constitute the Fast and



Slow Processes of Sensorimotor Learning. J Neurosci. 2015;35(26):9568-79.

31. Hinder MR, Tresilian JR, Riek S, Carson RG. The contribution of visual feedback to visuomotor adaptation: how much and when? Brain Res. 2008;1197:123-34.

32. Kawato M, Wolpert D. Internal models for motor control. Novartis Found Symp. 1998;218:291-304; discussion -7.