**CONTROLE NEURAL HIERÁRQUICO DE SELEÇÃO E APRENDIZAGEM DE AÇÕES**

**HIERARCHICAL NEURAL CONTROL OF ACTION SELECTION AND LEARNING**

Alan S. R. Fermin[[1]](#footnote-1)

**Resumo**

A realização de ações sequenciais é essencial para a sobrevivência humana e interações sociais. Apesar do progresso tecnológico para a gravação de atividade neural e achados recentes demonstrando como o cérebro controla e organiza ações sequenciais, ainda é necessário que se faça uma integração das abordagens teóricas e de evidências experimentais que melhor descrevem o controle neural do comportamento motor humano. Neste artigo, serão apresentados achados neurocientíficos recentes que ligam circuitos neurais organizados de forma paralela ao controle de ações sequenciais. Será proposto que esses circuitos neurais paralelos são, na realidade, controladores especializados na implementação de estratégias de seleção de ações e que essas estratégias assumem o controle do comportamento motor dependendo do contexto e do nível de experiência que indivíduos tem com certas tarefas. Serão mostradas evidências comportamentais e neurofuncionais da existência de três estratégias de seleção de ações (exploração, modelo-guiado, hábito-guiado) e como essas estratégias são realizadas por circuitos/controladores neurais conectando diferentes áreas do córtex frontal, dos gânglios de base e do cerebelo. Finalmente, será proposto uma extensão desse modelo paralelo de controle motor para um modelo neural hierárquico onde controladores diferentes trabalham de forma independente mas também enviam sinais entre si para auxiliar na arbitragem do controlador mais adequado para o controle motor dependendo do contexto.

**Palavras-chave**

Seleção de ações, circuitos córtico-subcorticais paralelos, controle hierárquico, inferência ativa

**Hierarchical Neural Control of Action Selection and Learning**

**Abstract**

Human survival and social interactions are highly dependent on the selection of sequential actions. Despite recent technological advances to record neural activity and recent findings demonstrating how the brain controls and organizes sequential actions, a broader theoretical approach is needed to integrate the theoretical advances and the experimental evidence that best explains the neural control of human sequential behavior. This opinion piece will discuss the recent neuroscientific findings that revealed the existence of parallel neural circuits implicated in the control of sequential actions. It will be argued that these parallel neural circuits are, in fact, neural controllers specialized at implementing action selection strategies and that such strategies assume behavior control in a context-dependent manner according to an individual’s experience. This opinion will also present behavioral evidence supporting the existence of three action selection strategies (exploratory, model-based, habit-based) and functional neuroimaging evidence demonstrating how these strategies are realized by distinct neural controllers linking the prefrontal cortex, the basal ganglia and the cerebellum. Finally, the parallel model of neural control of action selection will be extended into a hierarchical model where different controllers, represented by parallel neural circuits, work independently and also share information that supports controller arbitration.

**Key-words**

Action selection, cortico-subcortical circuits, hierarchical control, active inference

**INTRODUÇÃO**

Ações sequenciais são comportamentos observados em todos os seres vivos, como humanos, primatas, mamíferos, aves, animais aquáticos, insetos e até mesmo bactérias (ADAMS, 1984; BOLHUIS; OKANOYA; SCHARFF, 2010; FUJII; GRAYBIEL, 2003, 2003; JIN; COSTA, 2010; MAVRIDOU et al., 2018; THELEN, 1979). A seleção de ações é essencial não somente para a busca de meios que promovem a sobrevivência do indivíduo e manutenção das funções vitais do corpo (GENOVESIO; WISE; PASSINGHAM, 2014; KAMIL; KREBS; PULLIAM, 2012), mas também para processos cognitivos e comportamentais que são necessários para as relações sociais e que são afetados em desordens neurais (PERLSTEIN et al., 2001; RUITENBERG et al., 2015). A existência de comportamentos comuns à diversas espécies, principalmente entre os primatas, que compartilham quase 99% do seu código genetico com os humanos (MIKKELSEN et al., 2005; PRÜFER et al., 2012; SCALLY et al., 2012), sugere que possam existir mecanismos neurais comuns para o controle de ações em diversas espécies. O desenvolvimento de novas tecnologias para a observação e manipulação da atividade neural, como a ressonância magnética functional e a optogenética, assim como a utilização de modelos computacionais para explicar a função das áreas do cérebro, tem permitido entender como os seres vivos selecionam ações e quais os mecanismos neurais e respectivos processos cognitivos são empregados no comportamento motor de homens e outros animais.

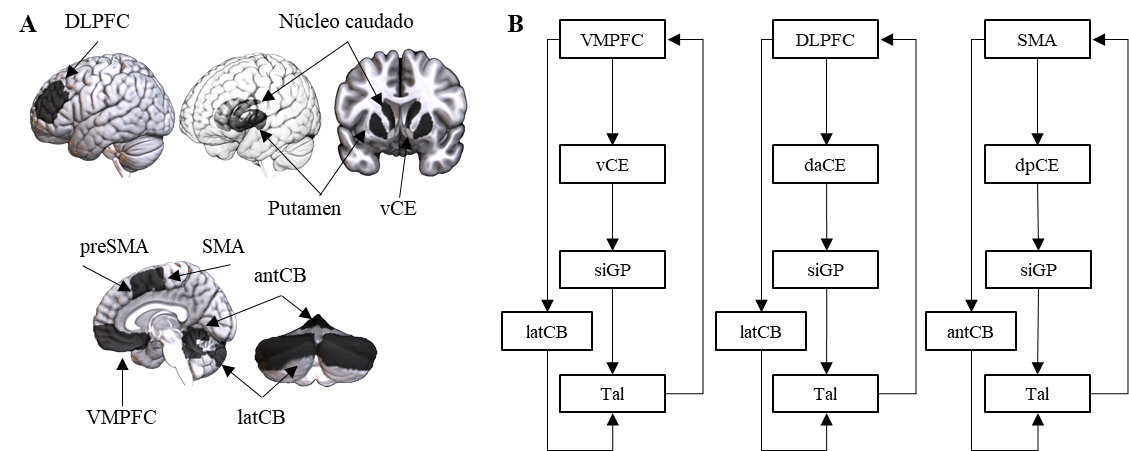
Nesse artigo de opinião, serão discutidas estratégias e modelos computacionais de seleção de ações. Essa opinião será focada em estudos com seres humanos e macacos, pois são os sujeitos que mais frequentemente participam em estudos de neurociência e que compartilham uma organização neural similar entre si. Primeiro será feita uma descrição da organização neuroanatômica de áreas do cérebro envolvidas no controle de ações sequenciais como o córtex prefrotal, córtex motor, os gânglios de base e o cerebelo. Na próxima seção, serão explicadas os modelos encontrados na literatura sobre estágios de aprendizagem de ações. Serão demonstradas evidências que associam estágios de aprendizagem à estratégias de seleção de ações e à circuitos neurais paralelos compostos for sub-áreas interligadas do cortex prefrontal, gânglios de base e cerebelo. Finalmente, será proposto um modelo hierárquico de controle neural para a seleção de ações.

**NEUROANATOMIA DO COMPORTAMENTO MOTOR**

Uma vez que os nomes de algumas áreas do cérebro serão citados na descrição das estratégias de seleção de ações, é preciso iniciar esse artigo com uma breve descrição dos sistemas neurais envolvidos no controle de movimentos. Será feita uma descrição das principais áreas cerebrais envolvidas no controle de ações sequenciais, mas o autor aconselha o leitor a procurar outros estudos envolvendo ações discretas como nas referências a seguir (SCHAAL et al., 2004; SHADMEHR; HOLCOMB, 1997; SHADMEHR; MUSSA-IVALDI, 1994; WOLPERT; GHAHRAMANI; JORDAN, 1995).

O Gráfico 1 mostra as áreas corticais frontais e as áreas subcorticais envolvidas na aprendizagem de comportamentos sequênciais (DOYON; PENHUNE; UNGERLEIDER, 2003; GRAFTON; HAZELTINE; IVRY, 1995; HIKOSAKA et al., 1999; JUEPTNER et al., 1997a, 1997b; LEHÉRICY et al., 2005; TONI et al., 1998). As quatro principais áreas do córtex frontal implicadas no sequenciamento de ações são: área ventro-medial prefrontal cortical (VMPFC), área dorso-lateral prefrontal cortical (DLPFC), área pre-suplementar motora (preSMA) e área suplementar motora (SMA). Os gânglios de base são um conjunto de vários núcleos celulares subcorticais. Para facilitar a discussão aqui pretendida, serão enfatizadas principalmente as funções da área ventral do corpo estriado (vCE) e da área dorsal do corpo estriado (dCE) dos gânglios de base. Em humanos, a área dCE é composta de dois núcleos, o núcleo caudado e o núcleo putamen, enquanto a área vCE é representada principalmente pelo núcleo acumbens. Ainda, a área dCE pode ser funcionalmente subdividida em área dorsal anterior (daCE) e dorsal posterior (dpCE). Uma outra área subcortical importante no controle dos movimentos sequenciais é o cerebelo (CB). Aqui, o foco serão as regiões anteriores (antCB) e lateral (latCB) do cerebelo.

As áreas descritas acima, entretanto, não funcionam isoladamente, pois estudos neuroanatômicos com animais e humanos demonstraram a existência de circuitos neurais paralelos ligando as áreas corticais às áreas subcorticais (ALEXANDER; DELONG; STRICK, 1986; DRAGANSKI et al., 2008; HABER, 2003; MIDDLETON; STRICK, 1994). A região VMPFC faz conexão anatômica predominantemente com as regiões vCE e latCB. A região DLPFC faz conexão anatômica com as regiões daCE e latCB. Finalmente, a área SMA faz conexão anatômica principalmente com as regiões dpCE e antCB. Apesar dos primeiros estudos neuranatômicos investigando as conexões entre as áreas corticais e subcorticas descritas acima sugerirem que esses circuitos funcionavam de forma segregada e com pouca comunicação entre si, estudos posteriores demonstraram que a existência de conexões laterais conectando esses circuitos neurais (HABER, 2003). De forma resumida, o circuito neural VMPFC-vCE-latCB faz conexão com o circuito DlPFC-daCE-latCB, que por sua vez faz conexão com o circuito SMA-dpCE-antCB. Essa organização anatômica com conexões laterais entre esses circuitos neurais sugere um possível controle hierárquico que será discutido adiante. Embora não representado no Gráfico 1, esses três circuitos também fazem conexões distintas com núcleos dos gânglios de base onde estão localizadas os neurônios dopaminérgicos (áreas ventral tegmentada e substância escura) que são importantes na aprendizagem guiada por sinais de recompensa (BROMBERG-MARTIN; MATSUMOTO; HIKOSAKA, 2010; SCHULTZ, 2006; SCHULTZ; DAYAN; MONTAGUE, 1997).



**Gráfico 1. Áreas do cérebro e circuitos neurais paralelos envolvidos no controle de ações sequenciais**. (**A**) Localização anatômica das principais áreas corticais e subcorticais envolvidas nas ações sequenciais. (**B**) Esquema gráfico dos circuitos neurais paralelos que ligam as áreas corticais prefrontais às áreas subcorticais dos gânglios de base e do cerebelo. Outros detalhes como conexões dentro dos gânglios de base e da organização quimíca e citológica das áreas vCE e dCE e dos núcleos do tálamo não estão representadas neste gráfico. Abreviações: DLPFC (dorso-lateral prefrontal córtex), VMPFC (ventro-medial prefrontal córtex), SMA (área suplementar motora, região posterior), pre-SMA (área suplementar motora, região anterior), antCB (cerebelo anterior), latCB (cerebelo lateral), vCE (corpo estriado ventral, a seta indica o núcleo acumbens), daCE (área dorsal anterior do corpo estriado), dpCE (área dorsal posterior do corpo estriado), siGP (seguimento interno do globo pálido), Tal (tálamus).

**ESTRATÉGIAS DE SELEÇÃO DE AÇÕES E MECANISMOS NEURAIS**

Para entender as funções dos circuitos neurais descritos acima é preciso entender as abordagens teóricas e os achados à nível de comportamento e neural associados às ações sequenciais. No estudo do comportamento motor, vários pesquisadores propuseram modelos de aprendizagem motora baseados na transição de estágios de aprendizagem (ADAMS, 1971; FITTS; POSNER, 1967; GENTILE, 1972; HIKOSAKA et al., 1999; NEWELL, 1985). Aqui, os modelos propostos por Fitts & Posner (FITTS; POSNER, 1967) e por Hikosaka e colaboradores (HIKOSAKA et al., 1999, 2002) serão apresentados em detalhes por conterem aspectos que são observados nos demais modelos. O modelo proposto for Fermin e colaboradores (FERMIN et al., 2010, 2016a) também será discutido por sugerir uma nova interpretação sobre os estágios de aprendizagem.

Fitts & Posner (FITTS; POSNER, 1967) propuseram um modelo teórico caracterizado por três estágios de aprendizagem. O estágio inicial, chamado de cognitivo-verbal, é caracterizado pelo alto uso de processos cognitivos, como a atenção e verbalização, visando conhecer as características da tarefa a ser aprendida, como seus objetivos, as regras e as informações sensoriais ligadas à tarefa (e.g. visual, auditiva). Como esse é um estágio inicial de aprendizagem, o comportanto é marcado por uma tomada de decisão mais lenta, uma maior duração na execuçao dos movimentos e um alto número de erros nas ações selecionadas. A transição para o estágio intermediário, chamado de associativo, se inicia quando os indivíduos demonstram um maior conhecimento das regras e informações da tarefa e eventualmente começam a utilizar esses conhecimentos para antecipar e associar os estímulos sensoriais à ações específicas para atingir os objetivos da tarefa. Portanto, o uso de processos cognitivos ainda são necessários nesse estágio, como o planejamento e antecipação de estímulos, porém, já é possível identificar uma redução no número de erros nas ações selecionadas. Finalmente, no estágio avançado, chamado de automático, os indivíduos já realizaram um alto número de tentativas e consequentemente atingem um alto índice de correção e rápida execução na seleção das ações. Essa rapidez na execução é geralmente interpretada como uma métrica de aquisição de automaticidade ou memória dos movimentos, quando os indivíduos já não precisam mais usar processos cognitivos nem informações exteroceptivas, mas somente selecionar ações guiadas por memórias das ações.

O próximo modelo de aprendizagem de ações a ser discutido em mais detalhes será aquele proposto por Okihide Hikosaka e collaboradores e desenvolvido por meio de estudos comportamentais e neurais com seres humans e macacos (HIKOSAKA et al., 1999). Esses estudos são provavelmente uns dos mais importantes da história da neurociência e foram realizados num período em que neurocientistas tentavam identificar as áreas do cérebro envolvidas na aprendizagem e armazenamento de memórias motoras e memórias declarativas (POLDRACK; PACKARD, 2003; SQUIRE, 1992, 2004; TULVING, 1987). De forma resumida, os primeiros estudos neurocientíficos investigando as bases neurais de diversos tipos de memórias revelaram que as funções de duas áreas do cérebro, o hipocampo e o córtex prefrontal, são essenciais para as memórias declarativas, enquanto as funções dos gânglios de base são essenciais para as memórias motoras. No entanto, como descrito na seção acima, estudos neuroanatômicos posteriores demonstraram uma organização neural mais complexa no processamento de informações (Gráfico 1) e esses achados influenciaram diretamente os trabalhos conduzidos por Hikosaka.

Os estudos neurocientíficos conduzidos por Hikosaka foram influenciados por teorias de aprendizagem e de controle mental (processo controlado e automático) (FITTS; POSNER, 1967; SHIFFRIN; SCHNEIDER, 1984). Hikosaka propôs uma hipótese integrativa que combinava partes dos processos cognitivos e neurais envolvidos nas memórias declarativas e motoras. Nessa hipótese, Hikosaka sugeriu que os processos cognitivos ligados à formação de memórias declarativas eram similares àqueles utilizados nos estágios iniciais de aprendizagem de ações, e que portanto, as mesmas áreas cerebrais ligadas à formação de memórias declarativas poderiam ser usadas nos estágios iniciais de aprendizagem de ações motoras. Já em relação às memórias motoras, Hikosaka sugeriu que essas poderiam ser armazenadas em áreas cerebrais distintas daquelas ligadas ao armazenamento de memórias declarativas. Aqui, aconselho o leitor a ler os artigos publicados por Hikosaka para conhecer os detalhes dos seus achados à nível de comportamento motor. No próximos parágrafos serão explicados seus achados que procuraram testar a hipótese descrita acima.

Para testar sua hipótese, Hikosaka criou a tarefa 2 x 5, baseada na associação de estímulos visuais e o aperto de botões usando os dedos da mãos. Imagine uma tela de computador, e no centro dessa tela aparece uma imagem de xadrez 4 x 4, isso significa quadrados de luzes organizados em quatro colunas e quatro linhas na tela, num total de 16 quadrados, sendo que cada quadrado podia ser iluminado de acordo com um programa de computador. No começo da tarefa, dois desses 16 quadrados são iluminados e os participantes, utilizando um pequeno controle de teclado com quatro botões (cada botão podendo ser apertado por um dedo específico - indicador, médio, anular e mindinho), tinham que descobrir os dois botões associados aos quadrados iluminados na tela e também a sequência para apertar esses botões. Se o participante acertasse a sequência dos dois primeiros botões, outros dois quadrados na tela eram iluminados e os participantes tinha que advinhar novamente a sequência desses novos botões. Essa tarefa se repetia num total de cinco combinações diferentes de dois quadrados iluminados na tela e portanto, ao final dessa tarefa, os participantes tinham que aprender uma sequência de dez apertos de botões, ou dez movimentos dos dedos, similar a uma sequência de notas de piano. Caso os participantes cometecem algum erro na sequênica de botões, as luzes da tela se apagavam e era necessário repetir a sequência do início até que os participantes conseguissem realizar a sequência de dez botões sem cometer erros.

Utilizando essa tarefa nos experimentos com humanos e com o uso de ressonância magnética de imagem funcional (fMRI), Hikosaka e seus colaboradores identificaram um aumento da atividade da área DLPFC e da área daCE no estágio inicial de aprendizagem quando os participantes cometiam um grande número de erros e tinham que utilizar processos cognitivos para entender e resolver a tarefa. Após inúmeras tentativas e já com poucos erros, a ativação da área DLPFC ainda estava presente. No entanto, quando os participantes já não mais cometiam erros e a velocidade de execução já era bem rápida, foi observada uma diminuição significativamente da ativação da área DLPFC e ao mesmo tempo um aumento na ativação de áreas motoras, como as áreas SMA e dpCE. Trabalhos realizados por outros pesquisadores, como os do grupo de Passingham, também observaram uma transição semelhante na ativação de áreas corticais e subcorticais, incluindo o cerebelo, durante a aprendizagem de movimentos sequenciais (DOYON; PENHUNE; UNGERLEIDER, 2003; JUEPTNER et al., 1997a, 1997b; LEHÉRICY et al., 2005; TONI et al., 1998).

Como o uso de fMRI apenas demonstra uma correlação entre atividade neural e comportamento, Hikosaka resolveu então provar uma relação causal entre função neural e comportamento (HIKOSAKA et al., 1999). Para isso, Hikosaka conduziu experimentos com macacos realizando cirurgias cerebrais para gravar a atividade neural e para a injeção de muscimol, um produto químico que aumenta temporariamente, lesão temporária, a atividade de neurônios inibidores e portanto, que leva à uma disfunção da área que recebe essa injeção. Esses estudos focaram nas áreas preSMA e SMA, daCE e dpCE, e em núcleos do cerebelo. A área preSMA tem conexão anatômica com a área DLPFC, portanto, manipulação de sua função reflete de forma indireta a função da área DLPFC. Esses estudos mostraram que: lesão temporária da área preSMA causou um alto número de erros na aprendizagem de novas ações sequências, enquanto lesão temporária na área SMA causou um número de erros similar em sequências novas e sequências já aprendidas. Resultados similares foram observados com lesões temporárias nos gânglios de base. Injeção de muscimol na área daCE levou a um aumento de erros nas novas sequências, enquanto injeção de muscimol na área dpCE levou a um aumento de erros nas sequências já aprendidas. Gravação da atividade neural também mostrou que neurônios na área preSMA mostraram uma maior ativação durante a aprendizagem de novas sequências comparado a ativação dos neurônios da área SMA. Esses estudos mostraram que as áreas DLPFC, preSMA e daCE são importantes no estágio inicial de aprendizagem de ações sequenciais, enquanto as áreas SMA e dpCE são importantes no estágio avançado. Ainda, como lesões temporárias nas áreas SMA e dpCE levaram a um aumento no número erros de sequências já aprendidas, esses resultados sugerem que essas áreas devem exercer alguma função no armazenamento de memórias motoras. Trabalhos recentes, no entanto, contestam esses achados e sugerem que uma das funções dos gânglios de base seja o controle de parâmetros dos movimentos, como tipo de contração, velocidade de contração, padrão espacial das ações, e não o armazenamento de memórias motoras (JURADO-PARRAS et al., 2020; LOPEZ-HUERTA et al., 2021; SUZUKI et al., 2021).

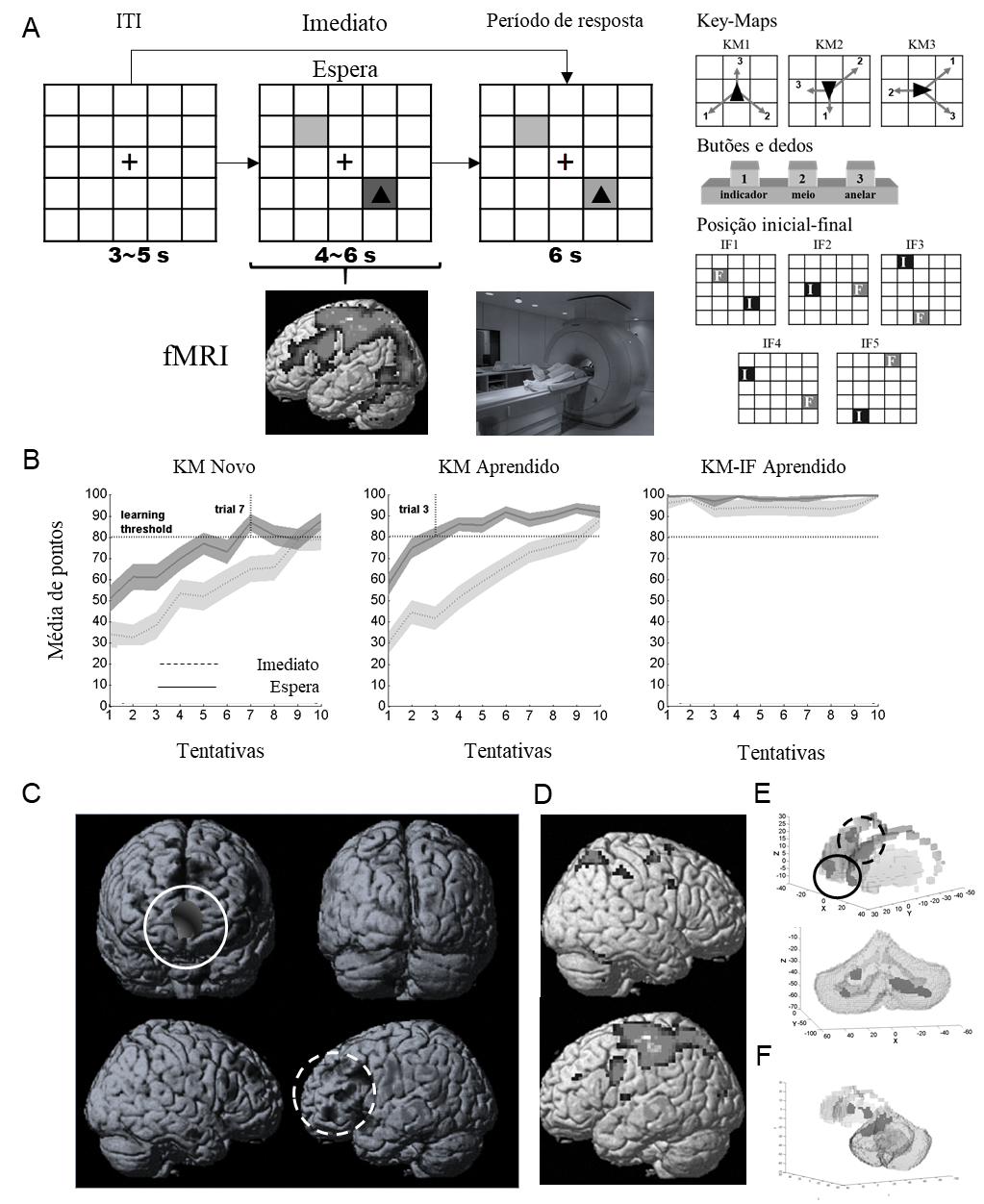
**CONTROLADORES NEURAIS INDEPENDENTES DE AÇÕES SEQUENCIAIS**

Os estudos neurocientíficos sobre os mecanismos neurais de aprendizagem e armazenamento de memórias de ações sequênciais descritos anteriormente, de uma forma geral, produziram evidências demonstrando a existência de áreas do cérebro e circuitos neurais cuja ativação aumenta ou diminui de acordo com a progressão do aprendizado e, portanto, suportam de certa forma as teorias de estágios de aprendizagem de ações. Duas perguntas, porém, emergem desses achados e teorias: (1) a transição entre estágios de aprendizagem é sempre sequencial? (2) áreas do cérebro ativadas nos estágios intermediário ou avançado de aprendizagem só podem ser ativadas se passarem pelo estágio que as precedem? Tentando resolver esse paradoxo, uma nova hipótese proposta por Fermin e colaboradores sugere que os estágios de aprendizagem na realidade refletem a seleção de estratégias de comportamentos sob o controle de múltiplos controladores neurais. De acordo com essa hipótese, cada controlador torna-se dominante no controle do comportamento dependendo do nível de experiência que indivíduos tem com certos contextos ambientais (FERMIN et al., 2010, 2016a).

A hipótese proposta por Fermin e colaboradores foi baseada na teoria de comportamento adaptivo da área da inteligência artificial conhecida como Reinforcement Learning (RL, termo orignal do inglês) (SUTTON; BARTO, 2018). Minha traduação livre para o português é Aprendizagem por Reforço. Aqui é importante fazer menção ao modelo de aprendizagem proposto por Adams (ADAMS, 1971) que, embora não elaborado em princípios matemáticos, segue uma idéia similar àquela proposta pela RL. A teoria RL sugere duas estratégias, ou controladores, de seleção e aprendizagem e de ações: modelo-livre (no inglês: model-free) e modelo-guiado (no inglês: model-based). A estratégia do modelo-livre funciona de forma semelhante às ações realizadas no estágio avançado de aprendizagem, ou seja, é utilizada para controlar o comportamento em contextos com os quais os agentes (e.g. robores, inteligência artificial, humanos ou animais) estejam bastante acostumados e nos quais as ações podem ser selecionadas de forma rápida, automática e sem a utilização de processes cognitivos complexos. Em contraste, a estratégia do modelo-guiado funciona com o uso de modelos internos, também chamados de modelos antecipatórios (no inglês: forward models). Modelos internos são, de forma geral, representações de situações, experiêcias, regras, personalidade dos outros, assim como o conhecimento de como máquinas funcionam, organização anatômica e fisiológica de seres vivos e até mesmo conhecimento acadêmico, como fórmulas matemáticas. A estratégia de modelo-guiado usa modelos internos para estimar a qualidade de ações hipotéticas e estimar a probabilidade de que a execução de uma determinada ação, se selecionada, pode levar o agente a atingir o objetivo da tarefa (DAW; NIV; DAYAN, 2005; DOYA, 1999; FERMIN et al., 2016a). Por exemplo, um jogador de futebol A, após assistir muitos jogos do jogador de futebol B, consegue entender que o jogador B sempre dribla para o lado esquerdo e, portanto, com base nesse modelo interno construído pela observação das habilidades do jogador B, o jogador A consegue fazer um julgamento e planejamento de quais ações ele precisa tomar para evitar que o jogador B execute o seu drible.

O modelo proposto por Fermin e colaboradores sugere a existência de três controladores neurais de comportamentos sequenciais: um controlador exploratório, um controlador de modelo-guiado e um controlador de hábitos ou comportamentos automáticos. O controlador exploratório, é utilizado em contextos com os quais os agentes não possuem qualquer experiência e nem conhecimento de suas variáveis e devem, portanto, explorar esses contextos para conhecer, por exemplo, as regras e os objetivos da tarefa e o conjunto de ações que podem ser executadas . O controlador de modelo-guiado, com explicado acima, faz o uso de modelos internos para estimar a qualidade e antecipar se uma determinação ação hipotética poder levar o agente a atingir o objetivo que deseja. Uma vez que o agente pode recorrer à esse modelo interno, o agente não mais necessita explorar o contexto e pode beneficiar-se com uma aprendizagem mais rápida. O controlador de hábito é caracterizado pelo armazenamento de mapas ou associações entre contextos e ações específicas. Esse mapas são adquiridos após inúmeras tentativas e repetições das ações que eventualmente se tornam automáticas e são selecionadas imediatamente quando o agente detecta que as características do contexto. A impressão que existe da transição de estágios se dá quando indivíduos precisam aprender uma tarefa, principalmente atividades mais complexas, e que exige a seleção de controladores diferentes com a progressão da aprendizagem.

Visando testar essa hipótese, Fermin e colaboradores realizaram uma série de experimentos comportamentais e com fMRI utilizando uma nova tarefa de navegação virtual (no inglês: grid-sailing task) que exige dos participantes realizar uma sequência de movimentos dos dedos para mover um cursor, localizado na tela de um computador, de sua posição inicial para uma posição final (Gráfico 2A). No começo da tarefa, uma imagem de xadrez, 5 x 5, é apresentada no centro da tela do computador. No centro da imagem do xadrez é exibida uma cruz para fixação visual para evitar que muitos movimentos dos olhos gerem sinais neurais que não possam ser interpretados na análise dos dados. Após um perído de 3 a 5 segundos, um cursor (e.g. triângulo) é apresentado na posição inicial juntamente com a posição final para a qual os participantes devem movimentar o cursor. A cor da posição inicial, no experimento original, podia ser vermelha ou verde. A cor vermelha indicava que os participantes tinham um período de 6 segundos para plenajar a sequência de ações, enquanto a cor verde indicava um início imediado das respostas (Gráfico 2A). Foram preparados três key-maps (KM, regras de movimentação do cursor) e cinco pares de posições iniciais e finais (IF). Para aprendizagem das ações sequenciais, os participantes foram instruídos que, se eles conseguissem atingir a posição final no menor número possível de ações, chamada de sequência perfeita, eles receberiam um feedback na tela indicado 100 pontos, e que para cada movimento adicional haveria um desconto de menos 5 pontos. Dessa forma, se a sequência perfeita fosse de 6 movimentos e o participante realizasse 8 movimentos, então o feedback exibido na tela seria 90 pontos, indicando aos particantes eles realizaram dois movimenots desnecessários.



**Gráfico 2. Tafefa de navegação e áreas cerebrais ativadas no planejamento da tarefa.** (**A**) Representação da organização da tarefa de navegação, as regras do KM e as posições IF. (**B**) Resultados do comportamento na realiação da tarefa nas condições Novo KM (condição 1), KM Aprendido (condição 2) e KF-IF Aprendido (condição 3). (**C, E**) Áreas cerebrais ativadas preferencialmente na condição 1 (círculo fechado) inclui VMPFC e vCE. Na condição 3 (círculo pontilhado) a ativação preferencial foi observada no DLPFC e dCE. (**D, F**) Áreas cerebrais ativas preferencialmente na condição 3 inclui o córtex motor e SMA e a parte posterior do núcleo putamen e antCB.

O experimento conduzido com fMRI foi realizado em dois dias: no primeiro dia para aprendizagem de algumas combinações de KM-IF, uma noite de descanso com o objetivo de consolidar a memória das ações sequenciais aprendidas no primeiro dia e, no segundo dia, foram feitas manipulações para investigar se diferentes condições experimentais levariam à ativações de sistemas neurais específicos ou não. Foram feitas três manipulações: condição 3 – repetir a sequência de ações aprendidas no primeiro dia; condição 2 – combinar os KMs aprendidos no primeiro dia com novas posições iniciais-finais; condição 1 – aprender um novo KM combinado com novas posiçõe iniciais-finais. De forma direta, era esperado na condição 1 que, como os participantes não tinham qualquer experiência com o novo KM, eles necessitassem explorar os movimentos para aprender a nova regra do KM. Já na condição 2, uma vez que os participantes já haviam aprendido as regras de certos KM (modelo ou regras do cursor), eles poderiam usar esse modelo interno para plenajar a sequência necessária para atingir a posição final. Ainda, era esperado que esse planejamento fosse melhor em tentativas quando havia um período de espera, pois os participantes poderiam usar esse tempo para antecipar e estimar a sequência de ações. Finalmente, na condição 3, uma vez que os participantes já haviam aprendido sequências de ações específicas para certas combinções entre KM-IF, foi esperado que eles não mais precisassem utilizar processos cogntivos elevados para selecionar as ações, mas somente utilizar as memórias das ações aprendidas no dia anterior.

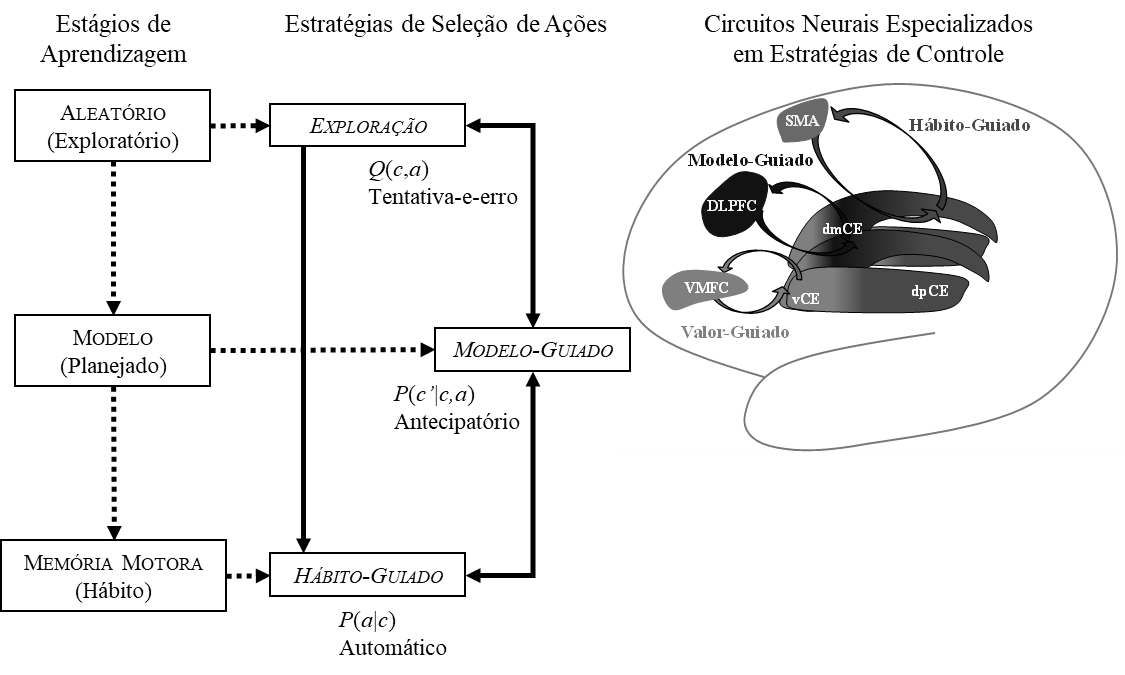
A série de estudos realizados por Fermin e colaboradores demonstrou que: na condição 1, foi observada uma aprendizagem mais lenta, com maior variabilidade e maior número de erros no aprendizado de novas sequências de ações com um novo KM, o que sugere que os participantes usaram uma estratégia exploratória (Gráfico 2B, painel da esquerda); na condição 2, uma vez que os participantes já tinham conhecimento do KM a ser utilizado, foi observado uma aprendizagem mais rápida, principalmente em tentativas com um tempo de espera quando o número de acertos foi mais alto do que em tentativas imediatas, o que sugere que os participantes utilizaram uma estratégia de modelo-guiado (Gráfico 2B, painel central); na condição 3, o número de acertos foi extremamente alto desde a primeira tentativa, o tempo de reação e a velocidade de execução também foram altamente rápidos em relação as condições 1 e 2, o que sugere a utilização de uma estratégia de hábito.

A análise da atividade neural gravada com fMRI focou no períod de espera com o objetivio de dissociar padrões de ativação neural relacionados a cada condição experimental. A análise demonstrou que: na condição 1, quando o comportamento foi mais exploratório, houve uma maior ativação nas áreas VMPFC e vCE (Gráfico 2C, E); na condição 2, quando o comportamento foi mais planejado, provavelmente pelo uso de uma estratégia de modelo-guiado, foi observado maior ativação nas áreas DLPFC, daCE e latCB (Gráfico 2C, E); finalmente, na condição 3, mesmo com o tempo de espera, houve uma maior ativação em áreas motoras como o córtex motor primário e SMA juntamente com ativação no dpCE e antCB (Grafico 2D, F).

Os achados neurais de outros estudos auxiliam na interpretação dada por Fermin e colaboradores de que os resultados do seu experimento de fMRI, descritos acima, são evidência de que o cérebro possui diferentes mecanismos neurais, ou controladores, que são especializados na implementação de estratégias de seleção de ações. Por exemplo, um experimento de fMRI identificou maior ativação na áreas VMPFC e vCE quando participantes decidiram explorar se nova ações poderiam gerar melhores resultados se comparadas com ações com as quais eles já eram familiares (DAW et al., 2006). Um outro estudo, utilizando uma tarefa de tomada de decisão observou maior ativação da área DLFPC quando participantes conseguiram utilizar com mais sucesso o conhecimento probabilístico de que a seleção de uma ação levaria à apresentação de estímulos visuais específicos como resultado da decisão (GLÄSCHER et al., 2010). Ativação da área DLPFC também é associada com o uso de modelos para a seleção de ações (DONOSO; COLLINS; KOECHLIN, 2014). Outros trabalhos também associam a ativação de áreas laterais do cerebelo com o uso de modelos internos (DOYA, 1999; IMAMIZU et al., 2000, 2003; WOLPERT; MIALL; KAWATO, 1998). Ativação das áreas SMA e dpCE também tem sido associadas com o uso de estratégia de hábito/automáticas (DOYON; PENHUNE; UNGERLEIDER, 2003; HIKOSAKA et al., 1999; WUNDERLICH; DAYAN; DOLAN, 2012).

**CARACTERÍSTICAS DO MODELO DE MÚLTIPLOS CONTROLADORES**

O Gráfico 3 apresenta uma representação esquemática do modelo de múltiplos controladores neurais de aprendizagem motora proposto por Fermin e colaboradores (FERMIN et al., 2010, 2016a). Do lado esquerdo, são representados três estágios de aprendizagem: exploratório, modelo-guiado e hábito-guiado. As setas com linhas quebradas ligando esses estágios tem o objetivo de sugerir que cada estágio pode ser executado de forma independente do estágio que o precede. Em contraste, na parte esquerda do modelo que apresenta os controladores ou estratégias de seleção de açõe são apresentadas as notações matemáticas utilizadas para caracterizar o método computacional utilizado por cada estratégia. É importante salientar que o processamento de feedback após a execução de ações faz parte do funcionamento de cada controlador neural.



**Gráfico 3. Modelo de múltiplos controladores neurais de comportamento**. Modelo proposto por Fermin e colaboradores. Ver texto para detalhes.

A estratégia de exploração, implementada pelo circuito neural VMPFC-vCE, é caracterizada pela variabilidade de comportamentos e por um alto uso de tentativas-e-erros. A notação *Q*(*c,a*) representa a aprendizagem da probilidade da qualidade, medida com métricas de feedback (e.g. alta pontuação, baixa pontuação), da ação (*a*) considerando o contexto (*c*) em que a ação é realizada. A estratégia de modelo-guiado, implementada pelo circuito neural DLPFC-daCE-latCB, é caracterizada pelo uso de modelos internos e antecipação dos resultados que poder ser alcançados por ações hipotéticas. A notação P(*c’*|*M*) representa o uso de um modelo interno de ações (*M*) com o objetivo de antecipar a probabilidade (*P*) de um futuro contexto (*c’*) que pode resultar em consequência da execução de uma determinada ação. Essa notação pode ser representada da forma P(*c’*|*c,a*) que decompõe o modelo (*M*) em uma combinação de contexte e ação (*c,a*). Finalmente, a estratégia de hábito ou comportamento automático, é caracterizada pela execução de ações previamente aprendizadas e que podem ser selecionadas de memória, sem o uso de informações exteriores. A estratégia de hábito é implementada pelo circuito neural SMA-dpCE-antCB. A notação P(*a*|*c*) representa a probabilidade da seleção de uma ação (*a*) considerando o contexto (*c*) identificado pelo agente.

As setas em linhas contínuas indicam que cada estratégia pode ser utilizada para auxiliar no funcionamento das outras estratégias e na progressão da aprendizagem. Por exemplo, muitas vezes a estratégia exploratória é selecionada para a aprendizagem e seleção de ações em novos contextos que requerem respostas bastante simples. Nesse caso, o indivíduo conseque aprender rapidamente as ações necessárias para completar a tarefa e o comportamento entre imediatamente sob o controle da estratégia de hábito. Portanto, a aprendizagem se torna mais rápida e sem requerer o uso de outra estratégia mais complexa, como a estratégia de modelo-guiado. Processo similar acontece com a estratégia de modelo-guiado, pois há momentos em que o indivíduo consegue identificar que pode utilizar modelos internos em outros contextos de aprendizagem e indicand,o dessa forma, não haver a necessidade de utilizar a estratégia exploratória. No entanto, também é possível que modelos internos possam ser utilizados inicialmente pela estratégia exploratória como forma de tentar associar informações previamente aprendidas com informações sendo adquiridas no novo contexto de aprendizagem. Da mesma forma, diferentes memórias motoras podem ser acessadas e usadas pela estratégia de modelo-guiado como forma de abstrair um modelo que possa ser utilizado para a seleção de ações.

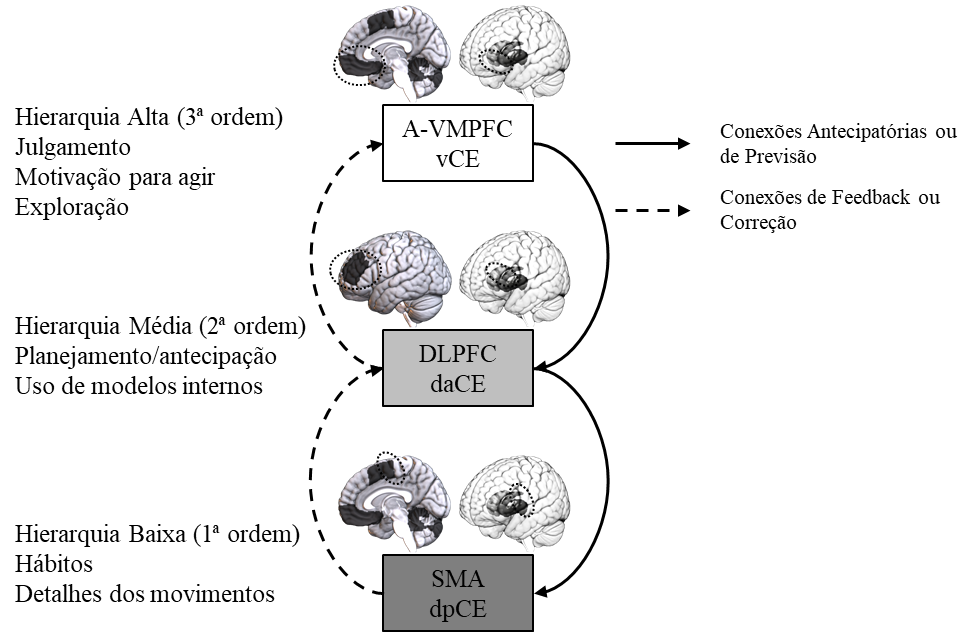
**ORGANIZAÇÃO HIERÁRQUICA DOS CONTROLADORES NEURAIS**

O modelo de múltiplos controladores neurais de seleção de ações proposto por Fermin e colaboradores foi baseado nos achados de estudos que investigaram as conexões neuroanatômicas que conectam áreas corticais à areas subcorticais. No presente artigo, foi dado um foco às conexões entre áreas do córtex prefrontal, gânglios de base e cerebelo, embora outros circuitos que ligam os córtices temporais e parietais à áreas suborticam também sejam importantes para o comportamento humano. No entanto, essa visão da organização paralela desses circuitos neurais não contempla todos os achados neuroanatômicos muitos menos as teorias mais recentes de funcionamento hierárquico do cérebro.

Como explicado na seção inicial desse artigo, estudos neuroanatômicos também apontam a existência de conexões laterais ligando circuitos neurais paralelos (BARBAS, 2015; BASTOS et al., 2012; HABER, 2003; PASSINGHAM; WISE, 2012; PETRIDES; PANDYA, 1999; STEMMLER; USHER; NIEBUR, 1995; ZEKI; SHIPP, 1988). Esses estudos sugerem uma relação bidirecional das conexões entre esses circuitos neurais, por exemplo, informação do córtex VMPFC é transmita para o córtex DLPFC. Da mesma forma, informação do córtex DLPFC é transmitida para o córtex SMA antes de chegar ao córtex motor primário que envia os sinais neurais para a coluna espinhal até alcançar os nervos motores. Informações também são enviadas de volta de uma área para outra, como por exemplo, do córtex DLPFC para o córtex VMPFC.

Com base nesses estudos neuroanatômicos e no modelo conhecido como Predictive Coding (minha tradução: Previsão Codificada) (HUANG; RAO, 2011; RAO; BALLARD, 1999), Karl Friston propôs um princípio de funcionamento do cérebro chamado de Free-Energy (minha tradução: Energia Livre) (FRISTON, 2009, 2010; FRISTON; KIEBEL, 2009; FRISTON; KILNER; HARRISON, 2006) que é utilizado num processo cognitivo chamado de Active Inference (minha tradução: Inferência Ativa) (ADAMS; SHIPP; FRISTON, 2013; FRISTON et al., 2013, 2016, 2017, 2017; FRISTON; MATTOUT; KILNER, 2011; FRISTON; SAMOTHRAKIS; MONTAGUE, 2012). Resumindamente, o princípio Free-Energy sugere que as regiões do cérebro armazenam modelos de informação que chegam de outras áreas. Ainda, as áreas do cérebro não só aprendem os padrões de informação vinda de outras áreas mas também aprendem a esperar o que essas áreas desejam. De acordo com o princípio Free-Energy, a função do cérebro é tentar diminuir a diferença entre a energia livre e a energia necessária para o seu funcionamento, em outras palavras, é diminuir a diferença entre o padrão de ativação esperado/previsto das outras áreas e o padrão de ativação que as áreas realmente assumem. Em outras palavras, todas as regiões do cérebro constroem modelos internos do padrão de ativação das áreas com as quais estão conectadas. O modelo Predictive Coding, baseado na existência de conexões laterais entre circuitos neurais ou entre áreas do cérebro, sugere que essas conexões paralelas representam, na realidade, uma organização hierárquica de transmissão de informações (BARDELLA et al., 2016; BASTOS et al., 2012; FRISTON, 2008; FRISTON; KIEBEL, 2009; HUANG; RAO, 2011; RAO; BALLARD, 1999). De acordo com esse modelo, as regiões cerebrais de maior hierarquia enviam sinais de previsão/antecipatórios, por meio de conexões antecipatórias, indicando o que espera da regiões de menor hierarquia. Por outro lado, as regiões de menor hierarquia enviam sinais de feedback/correção, através de conexões de feedback, para as regiões de maior hierarquia.

Combinando o princípio Free-Energy com o modelo Predictive Coding, Friston propôs a hipótese Active Inference para sugerir que sistemas neurais, incluindo o cérebro humano, procuram ativamente tentar diminuir o erro entre suas previões (e.g. resultados de ações, de estados do corpo humano e de mudanças no ambiente) e a informação real que chega ao cérebro. Portanto, o cérebro realiza inferências sobre estados futuros e estados reais de informação sensorial. A hipótese Active Inference junto com a organização hierárquica do cérebro, quando aplicadas aos circuitos neurais paralelos de ações sequênciais, sugere que o circuito SMA-dpCE-antCB é o circuito mais baixo na hierarquia, seguido pelo circuito intermediário composto por DLPFC-daCE-labCB e pelo circuito de mais alta hierarquia composto por VMPFC-vCE (Gráfico 4). Nessa organização hierárquica neural, o circuto de hábito SMA-dpCE-antCB deve ser o primeiro responsável pelo controle do comportamento através da identificação do contexto e da tarefa. Se há uma memória motora, essa memória será selecionada rapidamente para execução. No entanto, após um certo processamento computacional, se esse circuito de hábito se sente com pouca confidência e gera erros para controlar o comportamento, ele envia um signal de erro para o circuito imediatamente superior na hierarquia, o circuito de modelo-guiado composto por DLPFC-daCE-latCB. Computação semelhante é realizada pelo circuito de modelo-guiado que assume o controle do comportamento e recorre à modelos internos para antecipação dos resultados de ações hipotéticas. Se o circuito de modelo-guiado gera uma previsão com pouca confidência ou recebe um feedback de erro do movimento executado, esse envia um sinal de feedback para o circuito hierárquico imediatamente acima, o circuito de exploração VMPFC-vCE. Dessa forma, o circuito de exploração identifica que não existem memórias motoras nem modelos internos que possam ser utilizados para solucionar o problema atualmente enfrentado. Portanto, o circuito de exploração assume o controle do comportamento para realizar o conhecimento dessa nova tarefa. Outros modelos recentes também assumem uma organização hierárquica neural para o controle de movimentos tanto pelo córtex cerebral quando pelos gânglios de base (ITO; DOYA, 2011, 2015; PEZZULO; RIGOLI; FRISTON, 2015).



**Gráfico 4. Esquema hierárquica de múltiplos controladores neurais de comportamento**. Adaptação do modelo por Fermin e colaboradores baseado na organização hierárquica neuroanatômica e na hipótese Active Inference. Ver texto para detalhes.

**CONCLUSÃO E DIREÇÕES FUTURAS**

Nesse trabalho foram apresentados os achados neuroanatômicos que revelaram a organição de circuitos neurais paralelos e como esses circuitos estão envolvidos no controle de ações sequênciais. O modelo de múltiplos controladores neurais de comportamento proposto por Fermin e colaboradores foi descrito em detalhes por contemplar achados comportamentais e neurais que explicam com maior precisão o envolvimento de circuitos neurais paralelos na aprendizagem de novos comportamentos, no planejamento de ações através da utilização de modelos internos e na seleção de hábitos que são adquiridos com a formação de memórias motoras consolidadas após repetidas experiências. Também foi apresentada uma extensão hierárquica desse modelo para contemplar teorias mais recentes de funcionamento do cérebro. Estudos mais recentes procuram entender como esses sistemas neurais estão relacionados à comportamentos sociais (FERMIN et al., 2016b), à desordens de motivação, como observado em estados depressivos (BARRETT; QUIGLEY; HAMILTON, 2016; ICHIKAWA et al., 2020), desordens cognitivas observadas na esquizofrenia (ADAMS; PERRINET; FRISTON, 2012), desordens de controle motor observadas no Mal de Parkinson (REDGRAVE et al., 2010) e desordens desenvolvimentais como autismo (BAUMAN; KEMPER, 2005), dislexia (D’MELLO ANILA M.; GABRIELI JOHN D. E., 2018) e transtorno do desenvolvimento da coordenação (FUELSCHER et al., 2018). Também existe um esforço para entender como as conexões desses circuitos neurais são influenciadas por neuromoduladores como a serotonina, dopamina e acetilcolina (DAYAN, 2012; DOYA, 2008; DOYA; MIYAZAKI; MIYAZAKI, 2021; SCHULTZ; DAYAN; MONTAGUE, 1997). Futuros trabalhos são necessários para entender como o funcionamento e a interação entre múltiplos controladores neurais e as influências de neurmoduladores podem explicar diferenças individuais no aprendizado e controle de ações motoras ensinadas na Eduação Física.

**AGRADECIMENTOS**

Sou grato ao Professor Thomaz Abdalla pelo convite para escrever esse artigo. Também sou grato ao Ministério da Educação, Cultura, Esportes, Ciência e Tecnologia do governo Japonês pelo financiamento dos meus estudos.

**REFERÊNCIAS BIBLIOGRÁFICAS**

ADAMS, J. A. A Closed-Loop Theory of Motor Learning. **Journal of Motor Behavior**, v. 3, n. 2, p. 111–150, 1 jun. 1971.

ADAMS, J. A. Learning of movement sequences. **Psychological Bulletin**, v. 96, n. 1, p. 3, 1984.

ADAMS, R. A.; PERRINET, L. U.; FRISTON, K. Smooth Pursuit and Visual Occlusion: Active Inference and Oculomotor Control in Schizophrenia. **PLOS ONE**, v. 7, n. 10, p. e47502, 26 out. 2012.

ADAMS, R. A.; SHIPP, S.; FRISTON, K. J. Predictions not commands: active inference in the motor system. **Brain Structure and Function**, v. 218, n. 3, p. 611–643, maio 2013.

ALEXANDER, G. E.; DELONG, M. R.; STRICK, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. **Annual review of neuroscience**, v. 9, n. 1, p. 357–381, 1986.

BARBAS, H. General Cortical and Special Prefrontal Connections: Principles from Structure to Function. **Annual Review of Neuroscience**, v. 38, n. 1, p. 269–289, 2015.

BARDELLA, G. et al. Hierarchical organization of functional connectivity in the mouse brain: a complex network approach. **Scientific Reports**, v. 6, n. 1, p. 32060, 18 ago. 2016.

BARRETT, L. F.; QUIGLEY, K. S.; HAMILTON, P. An active inference theory of allostasis and interoception in depression. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 371, n. 1708, p. 20160011, 19 nov. 2016.

BASTOS, A. M. et al. Canonical Microcircuits for Predictive Coding. **Neuron**, v. 76, n. 4, p. 695–711, 21 nov. 2012.

BAUMAN, M. L.; KEMPER, T. L. Neuroanatomic observations of the brain in autism: a review and future directions. **International Journal of Developmental Neuroscience**, Autism: Modeling Human Brain Abnormalities in Developing Animal Systems. v. 23, n. 2, p. 183–187, 1 abr. 2005.

BOLHUIS, J. J.; OKANOYA, K.; SCHARFF, C. Twitter evolution: converging mechanisms in birdsong and human speech. **Nature Reviews Neuroscience**, v. 11, n. 11, p. 747–759, 2010.

BROMBERG-MARTIN, E. S.; MATSUMOTO, M.; HIKOSAKA, O. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. **Neuron**, v. 68, n. 5, p. 815–834, 9 dez. 2010.

DAW, N. D. et al. Cortical substrates for exploratory decisions in humans. **Nature**, v. 441, n. 7095, p. 876–879, jun. 2006.

DAW, N. D.; NIV, Y.; DAYAN, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. **Nature Neuroscience**, v. 8, n. 12, p. 1704–1711, dez. 2005.

DAYAN, P. Twenty-five lessons from computational neuromodulation. **Neuron**, v. 76, n. 1, p. 240–256, 2012.

D’MELLO ANILA M.; GABRIELI JOHN D. E. Cognitive Neuroscience of Dyslexia. **Language, Speech, and Hearing Services in Schools**, v. 49, n. 4, p. 798–809, 24 out. 2018.

DONOSO, M.; COLLINS, A. G. E.; KOECHLIN, E. Foundations of human reasoning in the prefrontal cortex. **Science**, v. 344, n. 6191, p. 1481–1486, 27 jun. 2014.

DOYA, K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? **Neural Networks**, v. 12, n. 7, p. 961–974, 1 out. 1999.

DOYA, K. Modulators of decision making. **Nature neuroscience**, v. 11, n. 4, p. 410–416, 2008.

DOYA, K.; MIYAZAKI, K. W.; MIYAZAKI, K. Serotonergic modulation of cognitive computations. **Current Opinion in Behavioral Sciences**, Computational cognitive neuroscience. v. 38, p. 116–123, 1 abr. 2021.

DOYON, J.; PENHUNE, V.; UNGERLEIDER, L. G. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. **Neuropsychologia**, v. 41, n. 3, p. 252–262, 2003.

DRAGANSKI, B. et al. Evidence for Segregated and Integrative Connectivity Patterns in the Human Basal Ganglia. **Journal of Neuroscience**, v. 28, n. 28, p. 7143–7152, 9 jul. 2008.

FERMIN, A. et al. Evidence for Model-Based Action Planning in a Sequential Finger Movement Task. **Journal of Motor Behavior**, v. 42, n. 6, p. 371–379, 29 out. 2010.

FERMIN, A. S. et al. Model-based action planning involves cortico-cerebellar and basal ganglia networks. **Scientific reports**, v. 6, p. 31378, 2016a.

FERMIN, A. S. et al. Representation of economic preferences in the structure and function of the amygdala and prefrontal cortex. **Scientific reports**, v. 6, p. 20982, 2016b.

FITTS, P. M.; POSNER, M. I. **Human performance**. Oxford, England: Brooks/Cole, 1967.

FRISTON, K. Hierarchical Models in the Brain. **PLOS Computational Biology**, v. 4, n. 11, p. e1000211, 7 nov. 2008.

FRISTON, K. The free-energy principle: a rough guide to the brain? **Trends in Cognitive Sciences**, v. 13, n. 7, p. 293–301, 1 jul. 2009.

FRISTON, K. The free-energy principle: a unified brain theory? **Nature Reviews Neuroscience**, v. 11, n. 2, p. 127–138, fev. 2010.

FRISTON, K. et al. The anatomy of choice: active inference and agency. **Frontiers in Human Neuroscience**, v. 7, 2013.

FRISTON, K. et al. Active inference and learning. **Neuroscience & Biobehavioral Reviews**, v. 68, p. 862–879, 1 set. 2016.

FRISTON, K. et al. Active Inference: A Process Theory. **Neural Computation**, v. 29, n. 1, p. 1–49, 1 jan. 2017.

FRISTON, K.; KIEBEL, S. Predictive coding under the free-energy principle. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 364, n. 1521, p. 1211–1221, 12 maio 2009.

FRISTON, K.; KILNER, J.; HARRISON, L. A free energy principle for the brain. **Journal of Physiology-Paris**, Theoretical and Computational Neuroscience: Understanding Brain Functions. v. 100, n. 1, p. 70–87, 1 jul. 2006.

FRISTON, K.; MATTOUT, J.; KILNER, J. Action understanding and active inference. **Biological Cybernetics**, v. 104, n. 1, p. 137–160, 1 fev. 2011.

FRISTON, K.; SAMOTHRAKIS, S.; MONTAGUE, R. Active inference and agency: optimal control without cost functions. **Biological Cybernetics**, v. 106, n. 8, p. 523–541, 1 out. 2012.

FUELSCHER, I. et al. Differential activation of brain areas in children with developmental coordination disorder during tasks of manual dexterity: An ALE meta-analysis. **Neuroscience & Biobehavioral Reviews**, v. 86, p. 77–84, 1 mar. 2018.

FUJII, N.; GRAYBIEL, A. M. Representation of action sequence boundaries by macaque prefrontal cortical neurons. **Science**, v. 301, n. 5637, p. 1246–1249, 2003.

GENOVESIO, A.; WISE, S. P.; PASSINGHAM, R. E. Prefrontal–parietal function: from foraging to foresight. **Trends in cognitive sciences**, v. 18, n. 2, p. 72–81, 2014.

GENTILE, A. M. A Working Model of Skill Acquisition with Application to Teaching. **Quest**, v. 17, n. 1, p. 3–23, 1 jan. 1972.

GLÄSCHER, J. et al. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. **Neuron**, v. 66, n. 4, p. 585–595, 27 maio 2010.

GRAFTON, S. T.; HAZELTINE, E.; IVRY, R. Functional mapping of sequence learning in normal humans. **Journal of cognitive neuroscience**, v. 7, n. 4, p. 497–510, 1995.

HABER, S. N. The primate basal ganglia: parallel and integrative networks. **Journal of Chemical Neuroanatomy**, Special Issue on the Human Brain - The Structural Basis for understanding Human Brain function and dysfunction. v. 26, n. 4, p. 317–330, 1 dez. 2003.

HIKOSAKA, O. et al. Parallel neural networks for learning sequential procedures. **Trends in neurosciences**, v. 22, n. 10, p. 464–471, 1999.

HIKOSAKA, O. et al. Central mechanisms of motor skill learning. **Current Opinion in Neurobiology**, v. 12, n. 2, p. 217–222, 1 abr. 2002.

HUANG, Y.; RAO, R. P. N. Predictive coding. **WIREs Cognitive Science**, v. 2, n. 5, p. 580–593, 2011.

ICHIKAWA, N. et al. Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. **Scientific reports**, v. 10, n. 1, p. 1–12, 2020.

IMAMIZU, H. et al. Human cerebellar activity reflecting an acquired internal model of a new tool. **Nature**, v. 403, n. 6766, p. 192–195, jan. 2000.

IMAMIZU, H. et al. Modular organization of internal models of tools in the human cerebellum. **Proceedings of the National Academy of Sciences**, v. 100, n. 9, p. 5461–5466, 29 abr. 2003.

ITO, M.; DOYA, K. Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. **Current Opinion in Neurobiology**, Behavioural and cognitive neuroscience. v. 21, n. 3, p. 368–373, 1 jun. 2011.

ITO, M.; DOYA, K. Distinct Neural Representation in the Dorsolateral, Dorsomedial, and Ventral Parts of the Striatum during Fixed- and Free-Choice Tasks. **Journal of Neuroscience**, v. 35, n. 8, p. 3499–3514, 25 fev. 2015.

JIN, X.; COSTA, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. **Nature**, v. 466, n. 7305, p. 457–462, 2010.

JUEPTNER, M. et al. Anatomy of motor learning. I. Frontal cortex and attention to action. **Journal of neurophysiology**, v. 77, n. 3, p. 1313–1324, 1997a.

JUEPTNER, M. et al. Anatomy of motor learning. II. Subcortical structures and learning by trial and error. **Journal of neurophysiology**, v. 77, n. 3, p. 1325–1337, 1997b.

JURADO-PARRAS, M.-T. et al. The Dorsal Striatum Energizes Motor Routines. **Current Biology**, v. 30, n. 22, p. 4362- 4372.e6, 16 nov. 2020.

KAMIL, A. C.; KREBS, J. R.; PULLIAM, H. R. **Foraging behavior**. [s.l.] Springer Science & Business Media, 2012.

LEHÉRICY, S. et al. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. **Proceedings of the National Academy of Sciences**, v. 102, n. 35, p. 12566–12571, 2005.

LOPEZ-HUERTA, V. G. et al. Striatal bilateral control of skilled forelimb movement. **Cell Reports**, v. 34, n. 3, p. 108651, 19 jan. 2021.

MAVRIDOU, D. A. et al. Bacteria use collective behavior to generate diverse combat strategies. **Current Biology**, v. 28, n. 3, p. 345–355, 2018.

MIDDLETON, F. A.; STRICK, P. L. Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. **Science**, v. 266, n. 5184, p. 458–461, 1994.

MIKKELSEN, T. et al. Initial sequence of the chimpanzee genome and comparison with the human genome. **Nature**, v. 437, n. 7055, p. 69–87, 2005.

NEWELL, K. M. Coordination, Control and Skill. In: GOODMAN, D.; WILBERG, R. B.; FRANKS, I. M. (Eds.). . **Advances in Psychology**. Differing Perspectives In Motor Learning, Memory, And Control. [s.l.] North-Holland, 1985. v. 27p. 295–317.

PASSINGHAM, R. E.; WISE, S. P. **The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight**. [s.l.] OUP Oxford, 2012.

PERLSTEIN, W. M. et al. Relation of prefrontal cortex dysfunction to working memory and symptoms in schizophrenia. **American Journal of Psychiatry**, v. 158, n. 7, p. 1105–1113, 2001.

PETRIDES, M.; PANDYA, D. N. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. **European Journal of Neuroscience**, v. 11, n. 3, p. 1011–1036, 1999.

PEZZULO, G.; RIGOLI, F.; FRISTON, K. Active Inference, homeostatic regulation and adaptive behavioural control. **Progress in Neurobiology**, v. 134, p. 17–35, 1 nov. 2015.

POLDRACK, R. A.; PACKARD, M. G. Competition among multiple memory systems: converging evidence from animal and human brain studies. **Neuropsychologia**, Functional Neuroimaging of Memory. v. 41, n. 3, p. 245–251, 1 jan. 2003.

PRÜFER, K. et al. The bonobo genome compared with the chimpanzee and human genomes. **Nature**, v. 486, n. 7404, p. 527–531, 2012.

RAO, R. P. N.; BALLARD, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. **Nature Neuroscience**, v. 2, n. 1, p. 79–87, jan. 1999.

REDGRAVE, P. et al. Goal-directed and habitual control in the basal ganglia: implications for Parkinson’s disease. **Nature Reviews Neuroscience**, v. 11, n. 11, p. 760–772, nov. 2010.

RUITENBERG, M. F. et al. Sequential movement skill in Parkinson’s disease: a state-of-the-art. **Cortex**, v. 65, p. 102–112, 2015.

SCALLY, A. et al. Insights into hominid evolution from the gorilla genome sequence. **Nature**, v. 483, n. 7388, p. 169–175, 2012.

SCHAAL, S. et al. Rhythmic arm movement is not discrete. **Nature neuroscience**, v. 7, n. 10, p. 1136–1143, 2004.

SCHULTZ, W. Behavioral Theories and the Neurophysiology of Reward. **Annual Review of Psychology**, v. 57, n. 1, p. 87–115, 2006.

SCHULTZ, W.; DAYAN, P.; MONTAGUE, P. R. A Neural Substrate of Prediction and Reward. **Science**, v. 275, n. 5306, p. 1593–1599, 14 mar. 1997.

SHADMEHR, R.; HOLCOMB, H. H. Neural correlates of motor memory consolidation. **Science**, v. 277, n. 5327, p. 821–825, 1997.

SHADMEHR, R.; MUSSA-IVALDI, F. A. Adaptive representation of dynamics during learning of a motor task. **Journal of neuroscience**, v. 14, n. 5, p. 3208–3224, 1994.

SHIFFRIN, R. M.; SCHNEIDER, W. Automatic and controlled processing revisited. **Psychological Review**, v. 91, n. 2, p. 269–276, 1984.

SQUIRE, L. R. Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. **Journal of Cognitive Neuroscience**, v. 4, n. 3, p. 232–243, 1 jul. 1992.

SQUIRE, L. R. Memory systems of the brain: A brief history and current perspective. **Neurobiology of Learning and Memory**, Multiple Memory Systems. v. 82, n. 3, p. 171–177, 1 nov. 2004.

STEMMLER, M.; USHER, M.; NIEBUR, E. Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics. **Science**, v. 269, n. 5232, p. 1877–1880, 29 set. 1995.

SUTTON, R. S.; BARTO, A. G. **Reinforcement Learning, second edition: An Introduction**. [s.l.] MIT Press, 2018.

SUZUKI, S. et al. Distinct regions of the striatum underlying effort, movement initiation and effort discounting. **Nature Human Behaviour**, v. 5, n. 3, p. 378–388, mar. 2021.

THELEN, E. Rhythmical stereotypies in normal human infants. **Animal behaviour**, v. 27, p. 699–715, 1979.

TONI, I. et al. The time course of changes during motor sequence learning: a whole-brain fMRI study. **Neuroimage**, v. 8, n. 1, p. 50–61, 1998.

TULVING, E. Multiple memory systems and consciousness. **Human Neurobiology**, v. 6, n. 2, p. 67–80, 1987.

WOLPERT, D. M.; GHAHRAMANI, Z.; JORDAN, M. I. An internal model for sensorimotor integration. **Science**, v. 269, n. 5232, p. 1880–1882, 1995.

WOLPERT, D. M.; MIALL, R. C.; KAWATO, M. Internal models in the cerebellum. **Trends in Cognitive Sciences**, v. 2, n. 9, p. 338–347, 1 set. 1998.

WUNDERLICH, K.; DAYAN, P.; DOLAN, R. J. Mapping value based planning and extensively trained choice in the human brain. **Nature Neuroscience**, v. 15, n. 5, p. 786–791, maio 2012.

ZEKI, S.; SHIPP, S. The functional logic of cortical connections. **Nature**, v. 335, n. 6188, p. 311–317, set. 1988.

1. Center for Brain, Mind and Kansei Sciences Research, Hiroshima University, Hiroshima, Japan Correspondência: [fermin@hiroshima-u.ac.jp](mailto:fermin@hiroshima-u.ac.jp) [↑](#footnote-ref-1)