

Mecanismos Cerebrais de Aprendizagem pelo Reforço na Perturbação Obsessivo-Compulsiva^{a)}

Brain Mechanisms of Reinforcement Learning in Obsessive-Compulsive Disorder

Pedro Castro-Rodrigues*, Bernardo Barahona-Corrêa**, Albino Oliveira-Maia**✉

RESUMO

Introdução: A Perturbação Obsessivo-Compulsiva (POC) é uma perturbação neuropsiquiátrica crónica cujo tratamento, farmacológico ou psicoterapêutico, é ineficaz numa percentagem significativa de casos, exigindo uma melhor compreensão da perturbação. Estudos realizados nestes doentes sugerem a existência de disfunção nos sistemas cerebrais responsáveis pela aprendizagem de novas acções, nomeadamente em circuitos cortico-estriados. No entanto, não há consenso sobre a forma como tal disfunção poderá estar relacionada com a sintomatologia obsessivo-compulsiva, nem sobre o potencial terapêutico de neuromodulação das áreas cerebrais envolvidas.

Objectivos: Rever os achados neuroimagiológicos e comportamentais relacionados com a aprendizagem de acções na POC bem como a sua eventual relevância terapêutica.

Métodos: Revisão não-sistemática da literatura, utilizando as palavras-chave “obsessive-compulsive disorder”, “neuroimaging”, “corticostriatothalamic dysfunction model”, “instrumental learning”, “reinforcement learning”, “deep brain stimulation” e “transcranial magnetic stimulation”.

Resultados: Em tarefas de aprendizagem pelo reforço, os doentes com POC apresentam dificuldade em desenvolver comportamentos intencionais ou baseados num modelo da tarefa, observando-se predominantemente comportamentos habituais. Por outro lado, os achados neuroimagiológicos mais consistentes na POC são a diminuição de volume e hiperactividade do córtex orbitofrontal (COF), assim como a hiperactividade do núcleo caudado, sendo que qualquer uma destas áreas tem envolvimento conhecido na aprendizagem pelo reforço. Na verdade, estudos em roedores permitiram

* Champalimaud Research and Clinical Centre, Champalimaud Centre for the Unknown, Lisboa; Centro Hospitalar Psiquiátrico de Lisboa, Lisboa; NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa.

** Champalimaud Research and Clinical Centre, Champalimaud Centre for the Unknown, Lisboa; Departamento de Psiquiatria e Saúde Mental, Centro Hospitalar de Lisboa Ocidental, Lisboa; NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa.

✉ albino.maia@neuro.fchampalimaud.org.

a) Baseado num trabalho apresentado oralmente no 7º Simpósio do Serviço de Psiquiatria do Hospital Professor Doutor Fernando Fonseca EPE, que teve lugar a 10 e 11 de Março de 2017.

demonstrar que o COF, em particular o COF lateral, desempenha um papel crucial no controlo da transição entre comportamentos habituais e comportamentos intencionais, sugerindo que este possa ser um défice nuclear na POC.

Conclusões: Apesar de todo o progresso nesta área, são necessárias novas tarefas de aprendizagem pelo reforço para caracterizar melhor o funcionamento do COF em doentes com POC, e para explorar os efeitos comportamentais da neuromodulação dessa região cerebral, com eventual potencial terapêutico.

Palavras-Chave: Perturbação Obsessivo-Compulsiva; Disfunção Cortico-Estriado-Talâmico-Cortical; Cortex Orbitofrontal; Aprendizagem pelo Reforço.

ABSTRACT

Introduction: *Obsessive-compulsive disorder is a chronic and highly prevalent neuropsychiatric disorder, with treatment lacking efficacy in a significant proportion of cases. Thus, a more comprehensive understanding of the brain mechanisms underlying OCD is needed. Dysfunction in corticostriatal circuits, which underlie learning of new actions, has been suggested by several studies. However, there is no consensus about how such dysfunction may generate obsessive-compulsive symptoms, nor about the therapeutic potential of invasive or non-invasive neuromodulation techniques targeting the brain areas that are involved.*

Objectives: *To review neuroimaging and behavioral findings related to learning of*

actions in OCD, and its potential therapeutic relevance.

Methods: *Non-systematic literature review, using the keywords “obsessive-compulsive disorder”, “neuroimaging”, “corticostriato-thalamic dysfunction model”, “instrumental learning”, “reinforcement learning”, “deep brain stimulation” e “transcranial magnetic stimulation”.*

Results: *In reinforcement learning tasks, OCD patients have a deficit in goal-directed or model-based learning, with a tendency towards habitual behavior or model-free learning. On the other hand, the most consistent neuroimaging findings in OCD are low volume and hyperactivity of the orbitofrontal cortex, as well as caudate hyperactivity – and both of these areas are important for reinforcement learning. In fact, rodent studies have demonstrated that the OFC, in particular the lateral OFC, plays an important role in the shift between habitual and goal-directed behavior, suggesting that this may be an important mechanism underlying OCD.*

Conclusions: *Notwithstanding all of the progress in this area, new reinforcement learning tasks are needed for improved understanding of the role of the OFC in OCD, and to explore the behavioral, and potentially also therapeutic effects of neuromodulation of this brain area.*

Key-Words: *Obsessive-Compulsive Disorder; Cortico-Striato-Thalamo-Cortical Dysfunction, Orbitofrontal Cortex, Reinforcement Learning.*

INTRODUÇÃO

A Perturbação Obsessivo-Compulsiva (POC) é uma perturbação neuropsiquiátrica crónica com prevalência estimada de 5.3% na população adulta portuguesa¹. É caracterizada pela presença de obsessões (ideias recorrentes, sentidas como intrusivas ou inapropriadas, e geradoras de ansiedade, angústia ou mau-estar) e/ou de compulsões (comportamentos ou actos mentais que o indivíduo afetado se sente compelido a fazer, normalmente de forma estereotipada ou ritualizada, e apesar de, frequentemente, reconhecer o seu carácter excessivo)^{2,3}. O tratamento da POC, baseado na combinação de farmacoterapia e psicoterapia, é eficaz em apenas 40 a 60% dos casos⁴⁻⁶, motivando estudos continuados para melhor compreender a natureza desta perturbação. Uma das áreas mais ativas de investigação tem como objectivo compreender de que forma a disfunção de circuitos cerebrais responsáveis pela aprendizagem e execução de ações poderá contribuir para a neurobiologia da POC. Na verdade, há evidência sugestiva da existência de um desequilíbrio entre os sistemas cerebrais responsáveis pela aprendizagem de ações^{7,8} entre os doentes com POC, nomeadamente com disfunção dos circuitos cortico-estriados⁹. No entanto não existe ainda consenso sobre a forma como tal possa estar relacionado com a sintomatologia obsessivo-compulsiva, nem sobre o valor terapêutico de neuromodulação das áreas cerebrais envolvidas^{10,11}.

OBJECTIVOS

Clarificar a relação entre achados neuroimiológicos e comportamentais no que diz respeito à aprendizagem de ações, na POC.

MATERIAL E MÉTODOS

Foi feita uma revisão não-sistemática da literatura, utilizando as palavras-chave “obsessive-compulsive disorder”, “neuroimaging”, “corticostriathalamic dysfunction model”, “instrumental learning”, “reinforcement learning”, “deep brain stimulation” e “cranial magnetic stimulation”.

RESULTADOS

Os primeiros estudos sobre aprendizagem de ações realizados em doentes com POC utilizaram tarefas inspiradas em paradigmas, dominantes na literatura animal, de condicionamento operante ou instrumental⁷. Estes paradigmas tinham anteriormente permitido, maioritariamente através de estudos em roedores utilizando recompensas alimentares para reforçar determinadas ações, distinguir comportamentos intencionais de comportamentos habituais^{12,13}. Assim, um comportamento é considerado intencional quando reflecte conhecimento da associação entre uma ação e as suas consequências e tem em conta o valor motivacional da ação^{13,14}. Em animais, previamente condicionados, tal poderá ser testado por manipulações da contingência entre a realização da ação e a obtenção da recompensa, ou do valor motivacional da recompensa (por exemplo, permitindo ao animal acesso livre à recompensa alimentar fora do contexto de condicionamento). Os

comportamentos intencionais são sensíveis a estas manipulações. Por outro lado, um comportamento pode tornar-se habitual através da repetição, ou da utilização de determinados paradigmas de condicionamento, resultando na existência de estímulos que dão origem a uma resposta comportamental de forma automática e independente das suas consequências^{13,14}. Assim, os comportamentos habituais não são sensíveis a manipulações da contingência ou do valor da recompensa (Figura 1).

O sistema intencional, sendo flexível, induz maior lentidão e consome maiores recursos cognitivos, enquanto o sistema habitual é rápido, mas à custa de menor flexibilidade na adaptação ao ambiente^{13,14}. Os circuitos neurais subjacentes a cada um destes sistemas foram já extensamente estudados, principalmente em roedores, com lesões bilaterais do estriado dorsomedial a dificultar a aprendizagem de ações intencionais¹⁵ e lesões do estriado dorsolateral a impedir a aquisição de hábitos¹⁶. Este tipo de paradigmas foi adaptado para uso em humanos¹⁷, tendo proporcionado achados em estudos de neuroimagem funcional que sugerem o envolvimento de áreas homólogas do estriado, nomeadamente no caudado e no putamen, respectivamente em ações de tipo intencional e habitual¹³. Há também evidência, recolhida principalmente em roedores, que permitiu demonstrar o envolvimento de outras áreas, tal como o córtex orbitofrontal (COF), neste tipo de comportamentos. Na verdade, o COF lateral desempenha um papel crucial no controlo da transição entre comportamentos habituais e comportamentos intencionais¹⁸.

A evidente proximidade entre hábitos, tal como descritos acima, e a fenomenologia da POC, assim como entre as áreas cerebrais envolvidas no condicionamento operante e as que estão alteradas na POC, motivaram a intersecção das duas áreas de estudo, tendo-se demonstrado que os pacientes com POC apresentam uma preponderância de comportamento operante do tipo habitual⁷. No entanto, tornou-se também claro que as tarefas de condicionamento operante, embora muito úteis para estudar comportamentos motivados por recompensas alimentares em roedores, são menos úteis para estudar comportamentos motivados por este e outros tipos de recompensas em humanos, encontrando-se dificuldades na sua replicação (Figura 2)¹⁴. Por outro lado, este tipo de tarefas necessitam de bastante treino antes da distinção entre diferentes tipos de comportamento poder ser testada, e geram poucos dados por cada indivíduo¹⁹. Assim, com o objectivo específico de estudar a aprendizagem das ações em humanos, foram desenvolvidos outros tipos de paradigmas, inspirados na teoria da aprendizagem pelo reforço (*reinforcement learning*)²⁰⁻²³.

A teoria da aprendizagem pelo reforço, que se originou na área da inteligência artificial, estuda a forma como qualquer tipo de sistema (biológico ou não biológico) pode aprender a escolher acções que maximizem a probabilidade de recompensa, e minimizem a probabilidade de punição²⁴. Segundo esta teoria, existe uma dicotomia entre um sistema de aprendizagem baseado num modelo (*model-based*) e um sistema de aprendizagem sem modelo (*model-free*)^{24,25}. Um agente que utilize o primeiro tipo de aprendizagem

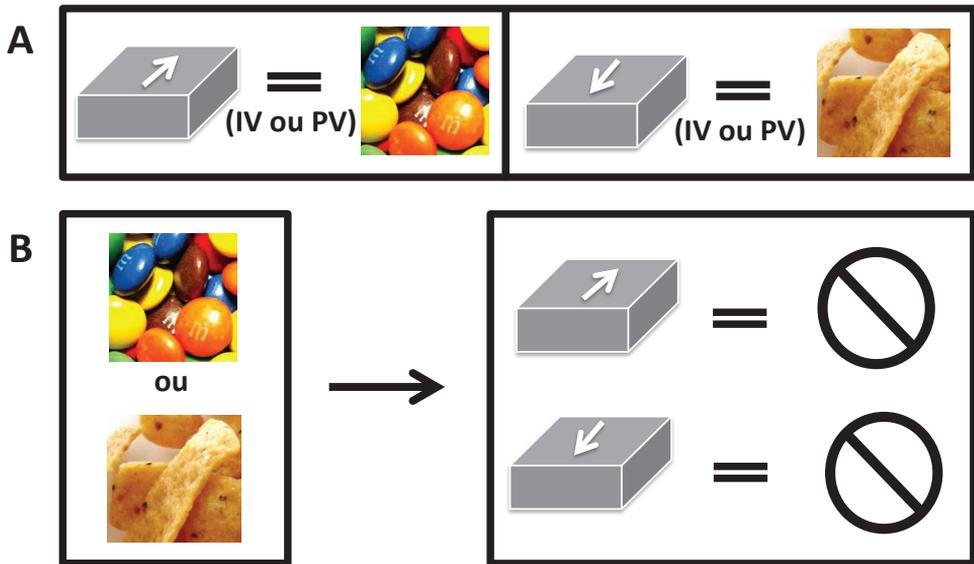


Figura 1 – Tarefa de Condicionamento Operante em Humanos.

Inspirados pelo trabalho de Tricomi e colegas (65) desenvolvemos uma tarefa de condicionamento operante que foi aplicada em voluntários saudáveis. A tarefa foi aplicada em laboratório com os voluntários em restrição alimentar. (A) Em cada sessão de condicionamento os voluntários tinham acesso a dois botões, cada um dos quais emparelhado com o acesso a uma recompensa alimentar (drageias de chocolate ou tiras de milho). As sessões tinham blocos consecutivos em que era dado acesso a uma ou à outra recompensa, dependendo do comportamento de manipulação do botão respectivo. Os blocos de cada recompensa, assim como blocos de descanso seguiam-se em ordem pseudo-randomizada, com indicação, em cada bloco, de qual botão estava ativo. Os participantes eram instruídos a carregar no botão tantas vezes quanto quisessem, sendo que a disponibilidade de uma recompensa era sinalizada com a imagem respectiva no monitor do computador. As recompensas eram atribuídas de forma probabilística de acordo com um de dois esquemas: um grupo de voluntários foi treinado num esquema de intervalo variável (IV) e outro grupo com um esquema de proporção variável (PV). No IV, a recompensa ficava disponível, em média, a cada 10 segundos, sendo atribuída na primeira vez que o botão fosse manipulado após o intervalo em causa para esse ensaio (IV-10). Na PV, a recompensa era atribuída, em média após o botão ser manipulado 30 vezes (PV-30). (B) Após períodos variáveis de treino, foi efectuada uma manipulação do valor motivacional da recompensa, para testar o grau de intencionalidade, vs. hábito, na manipulação do botão. Para esse efeito era permitido o acesso livre a uma das recompensas alimentares, para induzir a sua desvalorização por saciedade específica. Assim, uma das recompensas era desvalorizada, mas a outra não. Depois deste procedimento, o comportamento dos participantes no contexto experimental era novamente testado, por um período breve, e em extinção (ou seja, sem serem atribuídas recompensas) por forma a testar apenas a natureza do condicionamento prévio, e não o valor da recompensa propriamente dita. É de esperar que, quando o condicionamento induz o desenvolvimento de ações intencionais, tal como demonstrado em roedores para esquemas de treino de proporção variável (66), ocorram mais respostas para obter a recompensa valorizada do que a desvalorizada. Já para as ações habituais, que se espera resultarem de esquemas de treino de intervalo variável (66), não se espera que observem estas diferenças, uma vez que a execução da ação não depende já das suas consequências.

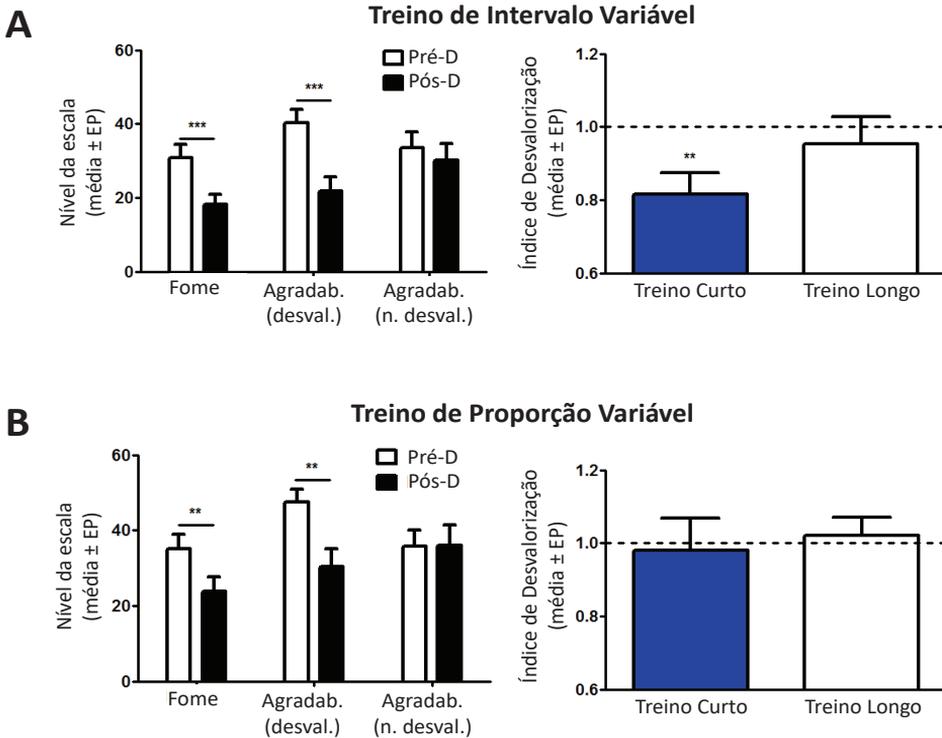


Figura 2 – Resultados de Condicionamento Operante em Voluntários Saudáveis.

(A) A tarefa de condicionamento descrita na figura 1 foi aplicada em 24 voluntários saudáveis com um esquema de intervalo variável (IV), 14 das quais com um protocolo de treino curto (2 sessões de 8 minutos no mesmo dia) e os restantes com um protocolo longo (3 sessões diárias de 8 minutos em 3 dias consecutivos). Após desvalorização de uma das recompensas, comparando valores pré-desvalorização (Pré-D) e pós-desvalorização (Pós-D), ocorreu, tal como seria de esperar, uma redução da fome, assim como da agradabilidade (agradab.) da recompensa desvalorizada (desvalor.), mas não da recompensa não desvalorizada (n. desvalor.). Foi também calculado um índice de desvalorização (IDesv), de acordo com a razão entre ações dirigidas para a recompensa desvalorizada e aquelas para obter a recompensa não desvalorizada, numa sessão efectuada imediatamente após a desvalorização. De acordo com trabalho prévio (65), a desvalorização teve impacto comportamental nos participantes com treino curto (IDesv < 1), mas não naqueles com treino longo (IDesv ≈ 1). Assim, o condicionamento com IV conduz inicialmente à aquisição de ações intencionais, que com o prolongamento do condicionamento transita para comportamento habitual. No entanto, este comportamento foi instável, não sendo muito claro na segunda metade de voluntários que foi testada. (B) Noutros 17 voluntários o condicionamento foi efectuado com um esquema de proporção variável (PV), 8 dos quais com treino curto e 9 com protocolo longo. Os achados relativos aos efeitos da desvalorização em escalas de fome e de agradabilidade foram sobreponíveis aos descritos acima. No entanto, ao contrário do que seria de esperar considerando dados recolhidos em modelo animal, o índice de desvalorização não indicou o desenvolvimento de comportamentos intencionais, independentemente da duração do condicionamento.

constrói um modelo interno do ambiente e aprende a relação entre os diferentes estados do mesmo, enquanto um agente que utilize o segundo tipo de aprendizagem funciona exclusivamente através da repetição de ações que levaram a recompensa no passado²⁴. As tarefas de aprendizagem pelo reforço, quando aplicadas em humanos, sustentam a existência dos dois tipos de ações, relacionados, tal como os comportamentos intencionais e habituais, com circuitos cerebrais distintos. No entanto, estas tarefas permitem gerar mais dados, com menos treino, do que os paradigmas de condicionamento operante¹⁹.

A tarefa de aprendizagem pelo reforço mais utilizada tem sido a *two-step task*, ou tarefa dos dois passos, em que o primeiro passo, de opção entre duas alternativas (I e II), serve para dar acesso a um de dois pares de alternativas (A e B ou C e D), para a decisão do segundo passo, que é aquela que vai determinar a obtenção de pequenas recompensas monetárias^{8,21,26}. Cada uma das opções do segundo passo (A, B, C, D) está associada a uma probabilidade de recompensa distinta e variável ao longo dos ensaios, pelo que a opção ideal na escolha do primeiro passo (I vs. II) será ter acesso ao par de opções de segundo passo (A e B ou C e D) que inclua a opção associada à maior probabilidade de recompensa naquele momento. No entanto, as transições entre opções do primeiro e do segundo passo não é fixa, mas antes probabilística, com transições comuns e transições raras (Figura 3). Assim, um agente com comportamento que esteja sustentado num modelo da tarefa (*model-based*, de tipo intencional) deverá manter a opção de primeiro

passo após transições comuns recompensadas e transições raras não recompensadas, e alterar a opção de primeiro passo após transições raras recompensadas e transições comuns não recompensadas. Neste tipo de tarefa, foi demonstrado que os doentes com POC têm dificuldade em desenvolver aprendizagem baseada num modelo da tarefa, ou seja, não parecem ser capazes de atuar de acordo com a previsão das consequências das suas opções, atuando sobretudo de acordo com a obtenção de recompensa no ensaio imediatamente anterior (sistema de aprendizagem sem modelo, *model-free*, de tipo habitual)²⁷. Curiosamente, na amostra de controlos saudáveis do mesmo estudo, foi encontrada uma correlação entre o volume do COF e o uso do sistema de aprendizagem baseado no modelo da tarefa²⁷.

Na verdade, os achados neuroimagiológicos mais consistentes na POC são precisamente a hiperactividade do COF e a sua diminuição de volume, bem como a hiperactividade do núcleo caudado^{9,28-30}. Por outro lado, um estudo recente de provocação de sintomas em humanos demonstrou que, durante um comportamento análogo a uma compulsão, existe diminuição da actividade do COF e aumento da actividade no putamen³¹. São estes achados que sustentam o modelo neurobiológico de disfunção cortico-estriada-talamo-cortical para a POC^{32,33}. Em paralelo, motivados em parte por achados de diminuição de sintomatologia obsessivo-compulsiva em doentes tratados com estimulação cerebral profunda (ECP, deep brain stimulation) por doença de Parkinson³⁴, foram realizados vários estudos para avaliar o potencial terapêu-

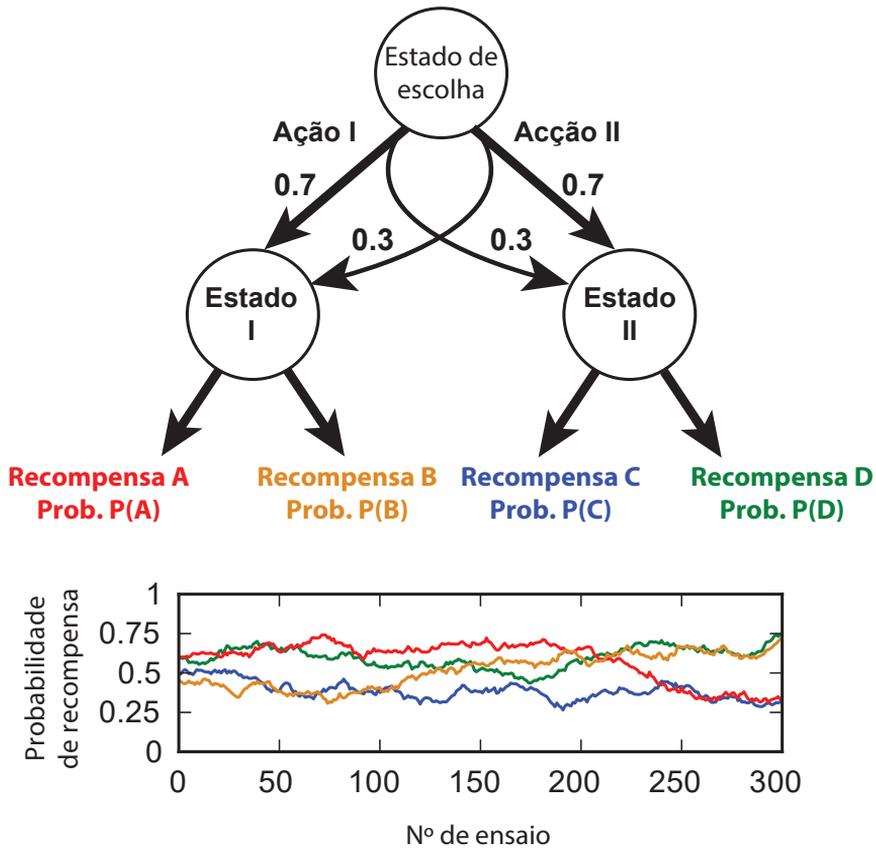


Figura 3 – Estrutura da Tarefa de Dois Passos.

A tarefa dos dois passos é a tarefa de aprendizagem pelo reforço mais utilizada em humanos. No primeiro passo o indivíduo opta entre duas alternativas (I e II), por forma a ter acesso a um outro par de alternativas entre as quais pode escolher (A e B ou C e D). No segundo passo, e dependendo do resultado do primeiro, o participante opta entre A e B ou entre C e D. É esta última opção que determina a obtenção de pequenas recompensas monetária, cujas probabilidades variam ao longo do tempo. Para poder escolher, no segundo passo, a opção com maior probabilidade de recompensa, a escolha do primeiro passo (I vs. II) deve ser aquela que dê acesso ao par de opções que a inclua. No entanto, as transições entre opções do primeiro e do segundo passo é também probabilística, com transições comuns ($P=0.7$) e transições raras ($P=0.3$). Adaptado de Akam et al, PLoS Computational Biology 2015. (67)

tico da ECP na POC. Estes estudos sugerem eficácia terapêutica de estimulação em cinco regiões cerebrais distintas, nomeadamente o braço anterior da cápsula interna^{35,36}, núcleo accumbens³⁷⁻³⁹, cápsula ventral/estriado ventral⁴⁰⁻⁴², núcleo subtalâmico^{43,44} e pedúnculo talâmico inferior⁴⁵. No entanto, não existe consenso sobre a região na qual a ECP é mais eficaz, nem existem factores preditores de resposta para cada doente¹⁰.

Há também estudos com recurso a estimulação magnética transcraniana (EMT, *transcranial magnetic stimulation*) para tratamento da POC, técnica que, permitindo neuromodulação focal de algumas zonas do córtex, tem a vantagem de ser não-invasiva. As áreas corticais para as quais estudos preliminares demonstraram eficácia terapêutica de neuromodulação por EMT foram o cortex prefrontal dorsolateral⁴⁶⁻⁴⁸, área motora suplementar⁴⁹, área motora pré-suplementar⁵⁰⁻⁵³ e córtex orbitofrontal^{54,55}. No entanto, há também estudos negativos para algumas destas áreas, nomeadamente o cortex prefrontal dorsolateral⁵⁶⁻⁵⁸ e áreas motoras suplementar⁵⁹ e pré-suplementar⁶⁰. Do nosso conhecimento, e segundo uma revisão publicada recentemente, não há, no entanto, estudos negativos com EMT do COF para tratamento da POC¹¹.

A variabilidade nos parâmetros dos estudos de neuromodulação terapêutica na POC, que não se esgota com os diferentes locais de estimulação, mas inclui também as alternativas nos parâmetros de estimulação (frequência e intensidade, entre outros), dificulta naturalmente a recolha de evidência nestas áreas. Por outro lado, é provável que o fenótipo cli-

nico da POC seja, em termos neurobiológicos, heterogéneo. A existência de um biomarcador que preveja uma resposta favorável a uma determinada intervenção terapêutica por neuromodulação seria fundamental para permitir o progresso para estas opções terapêuticas, e o comportamento no decorrer de tarefas de aprendizagem pelo reforço poderia ter esse papel no que diz respeito à neuromodulação dos circuitos implicados na aquisição desta aprendizagem. No entanto, não existe nenhum estudo no qual seja analisada a actividade cerebral em doentes com POC durante a realização de uma tarefa deste tipo e, tanto quanto nos foi possível apurar, não há dados que permitam avaliar os efeitos de neuromodulação, invasiva ou não-invasiva, na aprendizagem de ações em doentes com POC.

DISCUSSÃO E CONCLUSÕES

O desequilíbrio entre os sistemas cerebrais de aprendizagem de acções poderá estar na origem dos comportamentos repetitivos característicos da POC. No entanto, os paradigmas de aprendizagem pelo reforço desenvolvidos especificamente para uso em humanos têm algumas limitações. O paradigma mais popular, a tarefa dos dois passos, requer o uso de instruções complexas e alguns autores apontam este factor como problemático: os participantes não aprendem exclusiva e directamente a partir da sua interacção com a tarefa apresentada mas sobretudo através de informações verbais explícitas prévias⁶¹. Este factor dificulta a análise da evolução da aprendizagem com a experiência obtida exclusivamente a partir da interacção com

a tarefa. Para além disso, nesta tarefa o uso do sistema de aprendizagem baseado num modelo correlaciona-se com a memória de trabalho^{26,62}, o que pode dificultar a interpretação dos resultados em amostras clínicas nas quais a memória de trabalho possa estar afectada, tal como a POC⁶³. Assim são necessárias novas tarefas de aprendizagem pelo reforço que avaliem exclusivamente a aprendizagem a partir da tarefa e que permitam analisar a evolução da aprendizagem com a experiência.

Tendo em conta que, em modelos animais, o COF parece coordenar a transição entre comportamentos intencionais e habituais⁶⁴, e que, durante a provocação de sintomas em humanos, existe uma desativação do mesmo aquando da realização de um comportamento análogo a uma compulsão³¹, seria interessante explorar a actividade do COF em doentes com POC durante a realização de uma tarefa de aprendizagem pelo reforço. Para além disso, considerando que o COF é a única área do cérebro cujos estudos de estimulação magnética transcraniana apresentam exclusivamente resultados positivos, seria também interessante explorar os efeitos comportamentais (especificamente na aprendizagem de acções) da neuromodulação dessa área cerebral, verificando o seu potencial papel como biomarcador terapêutico em ensaios de neuromodulação.

Conflitos de Interesse / *Conflicting Interests:*

Os autores declaram não ter nenhum conflito de interesses relativamente ao presente artigo.

The authors have declared no competing interests exist.

Fontes de Financiamento / *Funding:*

O trabalho dos autores é suportado pelo projecto com a referência 30302, financiado por fundos nacionais através da Fundação para a Ciência e Tecnologia (FCT/MCTES) e co-financiado pelo FEDER no âmbito do Acordo de Parceria Lisboa 2020 – Programa Operacional Regional de Lisboa. A FCT/MCTES e o FEDER não tiveram qualquer papel no desenho e condução do estudo, incluindo na colheita, gestão, análise e interpretação dos dados, na preparação, revisão ou aprovação do manuscrito ou na decisão para submeter o manuscrito para publicação.

The authors are supported by grant 030302 funded by national funds from Fundação para a Ciência e Tecnologia (FCT/MCTES) and co-funded by FEDER, under the Partnership Agreement Lisboa 2020 - Programa Operacional Regional de Lisboa. FCT/MCTES and FEDER did not have a role in the design and conduct of the study, in the collection, management, analysis, and interpretation of the data, in the preparation, review, or approval of the manuscript, nor in the decision to submit the manuscript for publication.

Bibliografia / *References*

1. Caldas de Almeida J, Xavier M, Cardoso G, Pereira M, Gusmão R, Corrêa B, et al. Estudo Epidemiológico Nacional de Saúde Mental - 1.º Relatório. Lisboa Fac Ciências Médicas, da. 2013;60.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5). Diagnostic Stat Man Ment Disord 4th Ed TR. 2013;280.
3. Bernardo Barahona-Corrêa J, Camacho M, Castro-Rodrigues P, Costa R, Oliveira-Maia

- AJ. From thought to action: How the interplay between neuroscience and phenomenology changed our understanding of obsessive-compulsive disorder. Vol. 6, *Frontiers in Psychology*. 2015.
4. Hollander E, Bienstock CA, Koran LM, Pallanti S, Marazziti D, Rasmussen SA, et al. Refractory obsessive-compulsive disorder: State-of-the-art treatment. In: *Journal of Clinical Psychiatry*. 2002. p. 20–9.
 5. Kaplan A, Hollander E. A Review of Pharmacologic Treatments for Obsessive-Compulsive Disorder. *Psychiatr Serv* [Internet]. 2003;54(8):1111–8. Available from: <http://psychiatryonline.org/doi/abs/10.1176/appi.ps.54.8.1111>.
 6. Pallanti S, Quercioli L. Treatment-refractory obsessive-compulsive disorder: Methodological issues, operational definitions and therapeutic lines. *Prog Neuro-Psychopharmacology Biol Psychiatry*. 2006;30(3):400–12.
 7. Gillan CM, Pappmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW, et al. Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am J Psychiatry*. 2011;168(7):718–26.
 8. Voon V, Baek K, Enander J, Worbe Y, Morris LS, Harrison NA, et al. Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Transl Psychiatry*. 2015;5(11):e670.
 9. Menzies L, Chamberlain SR, Laird AR, Thelen SM, Sahakian BJ, Bullmore ET. Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: The orbitofronto-striatal model revisited. Vol. 32, *Neuroscience and Biobehavioral Reviews*. 2008. p. 525–49.
 10. Kohl S, Schönherr DM, Luigjes J, Denys D, Mueller UJ, Lenartz D, et al. Deep brain stimulation for treatment-refractory obsessive compulsive disorder: a systematic review. *BMC Psychiatry* [Internet]. 2014;14(1):214. Available from: <http://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-014-0214-y>.
 11. Zaman R, Robbins TW. Is there potential for Repetitive Transcranial Magnetic Stimulation (RTMS) as a treatment of OCD? In: *Psychiatria Danubina*. 2017. p. S672–8.
 12. Adams CD, Dickinson A. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol Sect B Comp Physiol Psychol*. 1981;33(2):109–21.
 13. Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010;35(1):48–69.
 14. Robbins TW, Costa RM. Habits. *Curr Biol* [Internet]. 2017;27(22):R1200–6. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0960982217312587>.
 15. Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci*. 2005;22(2):513–23.
 16. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*. 2004;19(1):181–9.
 17. de Wit S, Watson P, Harsay HA, Cohen MX, van de Vijver I, Ridderinkhof KR. Corticostriatal Connectivity Underlies Individual Differences

- in the Balance between Habitual and Goal-Directed Action Control. *J Neurosci* [Internet]. 2012;32(35):12066–75. Available from: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1088-12.2012>.
18. Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun*. 2013;4.
 19. Dolan RJ, Dayan P. Goals and habits in the brain. Vol. 80, *Neuron*. 2013. p. 312–25.
 20. Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010;66(4):585–95.
 21. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011;69(6):1204–15.
 22. Wunderlich K, Dayan P, Dolan RJ. Mapping value based planning and extensively trained choice in the human brain. *Nat Neurosci*. 2012;15(5):786–91.
 23. Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND. Model-based choices involve prospective neural activity. *Nat Neurosci*. 2015;18(5):767–72.
 24. Sutton RS, Barto AG. Introduction to Reinforcement Learning. *Learning*. 1998;4(1996):1–5.
 25. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005;8(12):1704–11.
 26. Smittenaar P, FitzGerald THB, Romei V, Wright ND, Dolan RJ. Disruption of Dorsolateral Prefrontal Cortex Decreases Model-Based in Favor of Model-free Control in Humans. *Neuron*. 2013;80(4):914–9.
 27. Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, et al. Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry*. 2015;20(3):345–52.
 28. Baxter Jr. LR, Phelps ME, Mazziotta JC, Guze BH, Schwartz JM, Selin CE. Local cerebral glucose metabolic rates in obsessive-compulsive disorder. A comparison with rates in unipolar depression and in normal controls. *Arch Gen Psychiatry* [Internet]. 1987;44(3):211–8. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3493749%5Chttp://archpsyc.jamanetwork.com/data/Journals/PSYCH/12426/archpsyc_44_3_003.pdf.
 29. Baxter LR, Schwartz JM, Mazziotta JC, Phelps ME, Pahl JJ, Guze BH, et al. Cerebral glucose metabolic rates in nondepressed patients with obsessive-compulsive disorder. *Am J Psychiatry*. 1988;145(12):1560–3.
 30. Nordahl TE, Benkelfat C, Semple WE, Gross M, King AC, Cohen RM. Cerebral glucose metabolic rates in obsessive compulsive disorder. *Neuropsychopharmacology*. 1989;2(1):23–8.
 31. Banca P, Voon V, Vestergaard MD, Philipiak G, Almeida I, Pociño F, et al. Imbalance in habitual versus goal directed neural systems during symptom provocation in obsessive-compulsive disorder. *Brain*. 2015;138(3):798–811.
 32. Saxena S, Rauch SL. Functional neuroimaging and the neuroanatomy of obsessive-compulsive disorder. Vol. 23, *Psychiatric Clinics of North America*. 2000. p. 563–86.
 33. Graybiel AM, Rauch SL. Toward a neurobiology of obsessive-compulsive disorder. *Neuron*. 2000;28(2):343–7.

34. Mallet L, Mesnage V, Houeto JL, Pelissolo A, Yelnik J, Behar C, et al. Compulsions, Parkinson's disease, and stimulation. *Lancet*. 2002;360(9342):1302–4.
35. Gabriëls L, Cosyns P, Nuttin B, Demeulemeester H, Gybels J. Deep brain stimulation for treatment-refractory obsessive-compulsive disorder: Psychopathological and neuropsychological outcome in three cases. *Acta Psychiatr Scand*. 2003;107(4):275–82.
36. Abelson JL, Curtis GC, Sagher O, Albucher RC, Harrigan M, Taylor SE, et al. Deep brain stimulation for refractory obsessive-compulsive disorder. *Biol Psychiatry*. 2005;57(5):510–6.
37. Denys D, Mantione M, Figeë M, van den Munkhof P, Koerselman F, Westenberg H, et al. Deep Brain Stimulation of the Nucleus Accumbens for Treatment-Refractory Obsessive-Compulsive Disorder. *Arch Gen Psychiatry* [Internet]. 2010;67(10):1061. Available from: <http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archgenpsychiatry.2010.122>.
38. Huff W, Lenartz D, Schormann M, Lee SH, Kuhn J, Koulousakis A, et al. Unilateral deep brain stimulation of the nucleus accumbens in patients with treatment-resistant obsessive-compulsive disorder: Outcomes after one year. *Clin Neurol Neurosurg*. 2010;112(2):137–43.
39. Roh D, Chang WS, Chang JW, Kim C-H. Long-term follow-up of deep brain stimulation for refractory obsessive-compulsive disorder. *Psychiatry Res* [Internet]. 2012;200(2–3):1067–70. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0165178112003290>.
40. Greenberg BD, Malone DA, Friehs GM, Reza AR, Kubu CS, Malloy PF, et al. Three-year outcomes in deep brain stimulation for highly resistant obsessive-compulsive disorder. *Neuropsychopharmacology*. 2006;31(11):2384–93.
41. Greenberg BD, Gabriëls LA, Malone DA, Reza AR, Friehs GM, Okun MS, et al. Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: Worldwide experience. *Mol Psychiatry*. 2010;15(1):64–79.
42. Tsai H-C, Chang C-H, Pan J-I, Hsieh H-J, Tsai S-T, Hung H-Y, et al. Pilot study of deep brain stimulation in refractory obsessive-compulsive disorder ethnic Chinese patients. *Psychiatry Clin Neurosci* [Internet]. 2012;66(4):303–12. Available from: <http://doi.wiley.com/10.1111/j.1440-1819.2012.02352.x>.
43. Mallet L, Polosan M, Jaafari N, Baup N, Welter M-L, Fontaine D, et al. Subthalamic nucleus stimulation in severe obsessive-compulsive disorder. *N Engl J Med* [Internet]. 2008;359(20):2121–34. Available from: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=19005196&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1056/NEJMoa0708514>.
44. Barcia JA, Reyes L, Arza R, Saceda J, Avecillas J, Yáñez R, et al. Deep brain stimulation for obsessive-compulsive disorder: Is the side relevant? *Stereotact Funct Neurosurg*. 2014;92(1):31–6.
45. Jiménez-Ponce F, Velasco-Campos F, Castro-Farfán G, Nicolini H, Velasco AL, Salín-Pascual R, et al. Preliminary study in patients with obsessive-compulsive disorder treated with electrical stimulation in the inferior thalamic peduncle. *Neurosurgery*. 2009;65(6 SUPPL. 1).
46. Greenberg BD, George MS, Martin JD, Benjamin J, Schlaepfer TE, Altemus M, et al. Effect of prefrontal repetitive transcranial magnetic stimulation in obsessive-compulsive di-

- sorder: a preliminary study. *Am J Psychiatry*. 1997;154(6):867–9.
47. Ma X, Huang Y, Liao L, Jin Y. A randomized double-blinded sham-controlled trial of electroencephalogram-guided transcranial magnetic stimulation for obsessive-compulsive disorder. *Chin Med J (Engl)* [Internet]. 2014;127(4):601–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24534207>.
48. Elbeh KAM, Elserogy YMB, Khalifa HE, Ahmed MA, Hafez MH, Khedr EM. Repetitive transcranial magnetic stimulation in the treatment of obsessive-compulsive disorders: Double blind randomized clinical trial. *Psychiatry Res*. 2016;238:264–9.
49. Mantovani A, Simpson HB, Fallon B a, Rossi S, Lisanby SH. Randomized sham-controlled trial of repetitive transcranial magnetic stimulation in treatment-resistant obsessive-compulsive disorder. *Int J Neuropsychopharmacol*. 2010;13(2):217–27.
50. Gomes P V, Brasil-Neto JP, Allam N, Rodrigues de Souza E. A randomized, double-blind trial of repetitive transcranial magnetic stimulation in obsessive-compulsive disorder with three-month follow-up. *J Neuropsychiatry Clin Neurosci*. 2012;24:437–43.
51. Mantovani A, Rossi S, Bassi BD, Simpson HB, Fallon BA, Lisanby SH. Modulation of motor cortex excitability in obsessive-compulsive disorder: An exploratory study on the relations of neurophysiology measures with clinical outcome. *Psychiatry Res*. 2013;210(3):1026–32.
52. Pallanti S, Marras A, Salerno L, Makris N, Hollander E. Better than treated as usual: Transcranial magnetic stimulation augmentation in selective serotonin reuptake inhibitor-refractory obsessive-compulsive disorder, mini-review and pilot open-label trial. *J Psychopharmacol*. 2016;30(6):568–78.
53. Lee Y-J, Koo B-H, Seo W-S, Kim H-G, Kim J-Y, Cheon E-J. Repetitive transcranial magnetic stimulation of the supplementary motor area in treatment-resistant obsessive-compulsive disorder: An open-label pilot study. *J Clin Neurosci Off J Neurosurg Soc Australas*. 2017;44:264–8.
54. Ruffini C, Locatelli M, Lucca A, Benedetti F, Insacco C, Smeraldi E. Augmentation effect of repetitive transcranial magnetic stimulation over the orbitofrontal cortex in drug-resistant obsessive-compulsive disorder patients: a controlled investigation. *Prim Care Companion J Clin Psychiatry* [Internet]. 2009;11(5):226–30. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2781034&tool=pmcentrez&rendertype=abstract>.
55. Nauczyciel C, Le Jeune F, Naudet F, Douabin S, Esquevin A, Vêrin M, et al. Repetitive transcranial magnetic stimulation over the orbitofrontal cortex for obsessive-compulsive disorder: A double-blind, crossover study. *Transl Psychiatry*. 2014;4(9).
56. Alonso P, Pujol J, Cardoner N, Benlloch L, Deus J, Menchón JM, et al. Right prefrontal repetitive transcranial magnetic stimulation in obsessive-compulsive disorder: a double-blind, placebo-controlled study. *Am J Psychiatry*. 2001;158(7):1143–5.
57. Sachdev PS, McBride R, Loo CK, Mitchell PB, Malhi GS, Croker VM. Right versus left prefrontal transcranial magnetic stimulation for obsessive-compulsive disorder: a preliminary investigation. *J Clin Psychiatry* [Internet]. 2001;62(12):981–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11780880>.

58. Sarkhel S, Sinha VK, Prahara SK. Adjunctive high-frequency right prefrontal repetitive transcranial magnetic stimulation (rTMS) was not effective in obsessive-compulsive disorder but improved secondary depression. *J Anxiety Disord.* 2010;24(5):535–9.
59. Kang JI, Kim C-H, Namkoong K, Lee C, Kim SJ. A Randomized Controlled Study of Sequentially Applied Repetitive Transcranial Magnetic Stimulation in Obsessive-Compulsive Disorder. *J Clin Psychiatry* [Internet]. 2009;70(12):1645–51. Available from: <http://article.psychiatrist.com/?ContentType=START&ID=10006172>.
60. Pelissolo A, Harika-Germaneau G, Rachid F, Gaudeau-Bosma C, Tanguy ML, BenAdhira R, et al. Repetitive transcranial magnetic stimulation to supplementary motor area in refractory obsessive-compulsive disorder treatment: A sham-controlled trial. *Int J Neuropsychopharmacol.* 2016;19(8):1–6.
61. Voon V, Reiter A, Sebold M, Groman S. Model-Based Control in Dimensional Psychiatry. Vol. 82, *Biological Psychiatry.* 2017. p. 391–400.
62. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci* [Internet]. 2013;110(52):20941–6. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1312011110>.
63. Abramovitch A, Abramowitz JS, Mittelman A. The neuropsychology of adult obsessive-compulsive disorder: A meta-analysis. Vol. 33, *Clinical Psychology Review.* 2013. p. 1163–71.
64. Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* [Internet]. 2013;4(May):2264. Available from: <http://www.nature.com/ncomms/2013/130806/ncomms3264/full/ncomms3264.html>.
65. Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci.* 2009;29(11):2225–32.
66. Hilario MRF. High on habits. *Front Neurosci* [Internet]. 2008;2(2):208–17. Available from: <http://journal.frontiersin.org/article/10.3389/neuro.01.030.2008/abstract>.
67. Akam T, Costa R, Dayan P. Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task. *PLoS Comput Biol.* 2015;11(12).