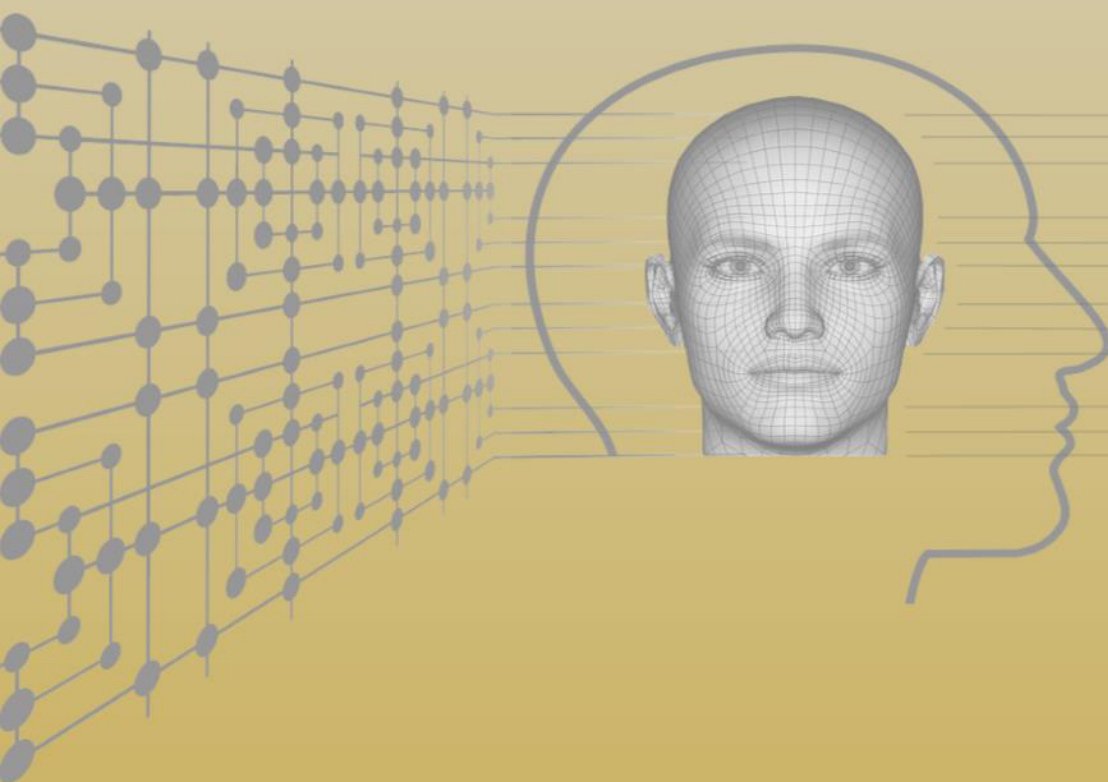


**PRÉMIO
CPC
CIÊNCIA**



**Miguel Santos Pais-Vieira
Carla Ferreira da Silva Pais-Vieira**

**PREVENÇÃO DA CORRUPÇÃO
E INFRAÇÕES CONEXAS
ASSOCIADAS À UTILIZAÇÃO DAS
INTERFACES CÉREBRO-MÁQUINA**

Título:

Prevenção da Corrupção e Infrações Conexas Associadas à Utilização das Interfaces Cérebro-Máquina

Autores:

Miguel Santos Pais-Vieira¹
Carla Ferreira da Silva Pais-Vieira¹

Série:

Cadernos CPC
Volume 1

Edição:

Conselho de Prevenção da Corrupção
(edição eletrónica | maio 2019)
ISBN: 978-989-20-9537-0

Contactos:

Av. da República, 65
1050-189 LISBOA
+351 217 945 138/9
cpc-tcontas@tcontas.pt
<http://www.cpc.tcontas.pt>

Apoios:

Fundação Bial: 96/1016; (MPV, CPV)
Fundação para a Ciência e a Tecnologia: FCT/IF/00098/2015 (MPV)

Copyright 2018/2019

A norma legal dos direitos de autor impede a reprodução por qualquer meio, sem prévio conhecimento e consentimento dos autores.

¹ Centro de Investigação Interdisciplinar em Saúde,
Instituto de Ciências da Saúde, Porto,
Universidade Católica Portuguesa

PREVENÇÃO DA CORRUPÇÃO
E INFRAÇÕES CONEXAS
ASSOCIADAS À UTILIZAÇÃO DAS
INTERFACES CÉREBRO-MÁQUINA

MIGUEL PAIS-VIEIRA | CARLA PAIS-VIEIRA

CADERNOS DO CPC

O Conselho de Prevenção da Corrupção criou o *Prémio CPC-Ciência* por Deliberação de 04 de maio de 2017. A iniciativa teve em conta quer o interesse público em promover o estudo e a pesquisa das matérias relacionadas com a sua missão, quer o diálogo institucional e a cooperação com universidades e centros de investigação na definição de projetos comuns para aprofundar o conhecimento do fenómeno corruptivo e a procura de concretas soluções preventivas.

A fim de divulgar as investigações vencedoras do galardão, nos termos do seu Regulamento, surge esta edição dos Cadernos do CPC. Assim o Conselho visa contribuir para a disseminação junto do grande público do que melhor a Academia portuguesa vai produzindo nos domínios da ética, integridade, transparência e gestão dos riscos de corrupção em todas as suas formas. Nesta categoria se insere o trabalho dos investigadores Miguel Pais-Vieira e Carla Pais-Vieira, premiados da edição de 2018.

RESUMO

A integridade, a responsabilização e a transparência são fatores cruciais para evitar atos de corrupção passiva ou ativa. As novas tecnologias associadas à manipulação e utilização de dados com origem na atividade cerebral – as interfaces cérebro-máquina – apresentam um potencial de negócio de 1,24 mil milhões de euros até 2020 e são particularmente suscetíveis à intrusão (*‘hacking’*) por terceiros. Estas intrusões, a possibilidade de vantagem indevida, e os abusos de poder daí decorrentes, podem ocorrer através da extração de dados relevantes ou através da manipulação de zonas cerebrais que levem os utilizadores a realizar tarefas sem o seu conhecimento ou consentimento. O presente estudo integra as áreas do Direito, da Saúde Pública e da Bioética e visa: 1) descrever os princípios através dos quais operam as interfaces cérebro-máquina, 2) referir de que forma estes princípios podem ser alvo de situações de abuso de poder por parte de terceiros, e 3) estabelecer medidas objetivas para a redução do risco de situações potenciadoras de vantagem indevida e abuso de poder associados a estas tecnologias.

Palavras-chave: sentido de agência, vantagem indevida, abuso de poder, interfaces cérebro-máquina, neuromodulação, motivação

ABSTRACT

Integrity, accountability, and transparency are key factors to consider when avoiding passive or active corruption. The new technologies associated with brain manipulation or with the use of brain activity – brain-machine interfaces – have a potential for business estimated in 1.24 billion euros up to 2020. These technologies are particularly susceptible to hacking by third parties because brain manipulation through such devices may allow extracting information or biasing the user to perform a particular action without it’s knowledge or consent, therefore potentiating unfair advantage and abuse of power. The present study is relevant for the fields of Law and Public Health and is aimed at: 1) describing the main components and variables interfering with brain-machine interfaces operation, 2) describe how each of these components may be subject to abuse of power by third parties, and 3) make objective recommendations to reduce situations associated with unfair advantage and abuse of power when using brain-machine interfaces.

Keywords: sense of agency, unfair advantage, abuse of power, brain-machine interfaces, neuromodulation, motivation

ÍNDICE GERAL

PREÂMBULO	8
1 – INTRODUÇÃO	9
2 – RELAÇÃO ENTRE INTERFACES CÉREBRO-MÁQUINA E CORRUPÇÃO	12
3 – DEFINIÇÃO DE INTERFACES CÉREBRO-MÁQUINA	13
3.1 – COMPONENTES GERAIS.....	16
4 – O SENTIDO DE AGÊNCIA	17
5 – COMPONENTES ESPECÍFICOS DAS INTERFACES CÉREBRO-MÁQUINA QUE PODEM INTERFERIR COM A RESPONSABILIDADE	21
5.1 – NODO E NÚMERO DE NODOS	21
5.2 – ARQUITETURA DA REDE NEURONAL	24
5.3 – CARACTERÍSTICAS DA TÉCNICA DE REGISTO, DE DESCODIFICAÇÃO, E DE MODULAÇÃO.....	26
5.4 – AÇÃO ITERATIVA OU CONTÍNUA	29
5.5 – NATUREZA DA TAREFA A REALIZAR	32
5.6 – ENVOLVIMENTO DO SUJEITO	33
5.7 – RECOMPENSA	37
5.8 – VOLIÇÃO	40
6 – NOTAS FINAIS	41
7 – CONCLUSÕES	43
REFERÊNCIAS	44

PREÂMBULO

O presente texto visa descrever os princípios que estão envolvidos na utilização das interfaces cérebro-máquina, por forma a permitir a elaboração de um enquadramento legal que possa prevenir situações de corrupção derivadas da má utilização desta nova tecnologia. Os autores encontram-se envolvidos no estudo dos mecanismos neurofisiológicos da mente humana, bem como no desenvolvimento de interfaces cérebro-máquina. É intenção dos autores apresentar um texto simples que permita a académicos e não-académicos uma compreensão de quais as variáveis que, estando envolvidas nas interfaces cérebro-máquina e no funcionamento do cérebro humano, podem condicionar o sentido de agência, a tomada de decisão e, conseqüentemente, a responsabilidade legal que o indivíduo apresenta sobre as suas ações. Da mesma forma, pretende-se estabelecer um conjunto de princípios que podem contribuir para a prevenção de situações de abuso através da intrusão por terceiros ('hacking'), que possam resultar na apropriação e utilização inadvertida de dados do utilizador, na indução de computação associada à atividade cerebral dos utilizadores ou mesmo na indução à realização de ações específicas.

Veremos que a manipulação dos circuitos cerebrais e o registo da atividade dos mesmos pode levar à realização de determinadas ações sem que o indivíduo tenha um pleno estado de consciência do móbil dos seus atos. A prevenção da corrupção durante a utilização das interfaces cérebro-máquina deverá ter em conta o fabricante, o clínico e o utilizador. Estes agentes podem interferir com diferentes variáveis envolvidas na interface cérebro-máquina. Torna-se assim necessário fazer uma descrição técnica simples, mas suficientemente detalhada, e uma descrição da forma como estas variáveis e intervenientes interagem, de modo a permitir a elaboração de leis adequadas a esta tecnologia.

Em suma, as tecnologias que utilizam a atividade cerebral para controlo de máquinas apresentam um enorme potencial para o restabelecimento e expansão de funções do corpo humano; mas também apresentam inúmeras fragilidades para as quais não existe ainda um enquadramento legal detalhado. A delimitação de princípios e variáveis necessários para o controlo adequado desta tecnologia dirigida a não-peritos, constitui assim uma prioridade, de modo a que se possa estabelecer o enquadramento legal apropriado.

1 – INTRODUÇÃO

O cérebro humano saudável processa informação e tem a capacidade de enviar ordens para os diferentes sistemas corporais a fim de manter o funcionamento normal do corpo. Por exemplo, quando o sujeito deseja mover o seu braço, há um comando vindo do cérebro que depois de percorrer vias específicas do sistema nervoso e músculos, resulta num movimento final (ver Figura 1 painel à esquerda). Em situações de lesão, a interrupção desta comunicação entre o cérebro e os diferentes componentes do corpo (por exemplo: ossos, músculos, etc.) impede o normal funcionamento do corpo, tal como acontece em situações de lesão vertebro-medular (ver Figura 1, painel à direita).

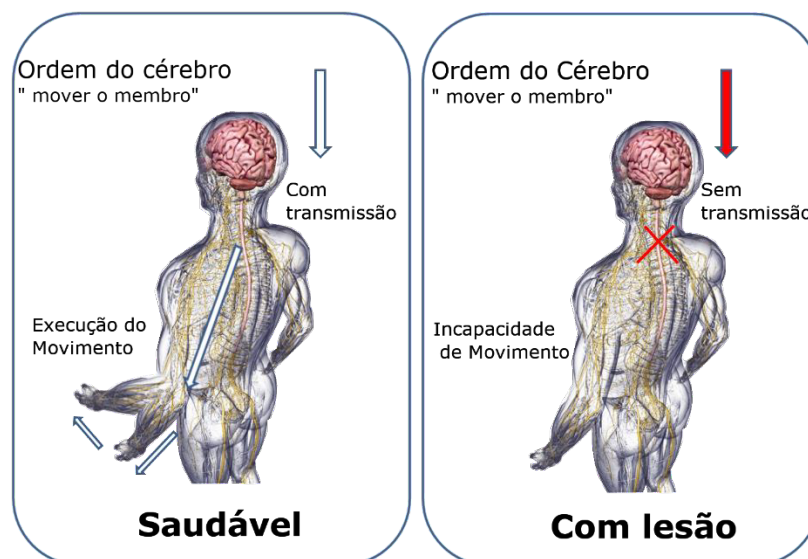


Figura 1 – Transmissão de informação no ser humano saudável e com lesão vertebro-medular.

As interfaces cérebro-máquina são engenhos que utilizam a atividade do sistema nervoso para controlar atuadores de forma a permitir a realização de funções comportamentais (isto é: realização de múltiplas tarefas) (Lebedev e Nicolelis, 2017). De uma forma não técnica, poder-se-ia descrever uma interface cérebro-máquina como um mecanismo que controla ou é controlado pela atividade do cérebro¹. A principal motivação que levou ao desenvolvimento das interfaces cérebro-máquina prendia-se com a possibilidade de, após a lesão de uma parte do corpo, poder utilizar a informação

¹ Veremos adiante que os interfaces cérebro-máquina podem envolver outros componentes do sistema nervoso.

que ainda se mantém intacta no sistema nervoso para controlar uma máquina que, por sua vez, iria realizar a função desejada. Por outras palavras, concluiu-se que o facto de um sujeito apresentar um défice, por exemplo no membro superior, não implicava necessariamente que houvesse dano nas estruturas do sistema nervoso que programam e decidem a utilização desse membro (Figura 2 painel à esquerda). Assim, a utilização de estruturas do sistema nervoso que se encontram intactas para gerar comandos, os quais posteriormente são executados por uma máquina constitui o racional terapêutico das interfaces cérebro-máquina (por exemplo, um braço robótico, Figura 2 painel à direita).

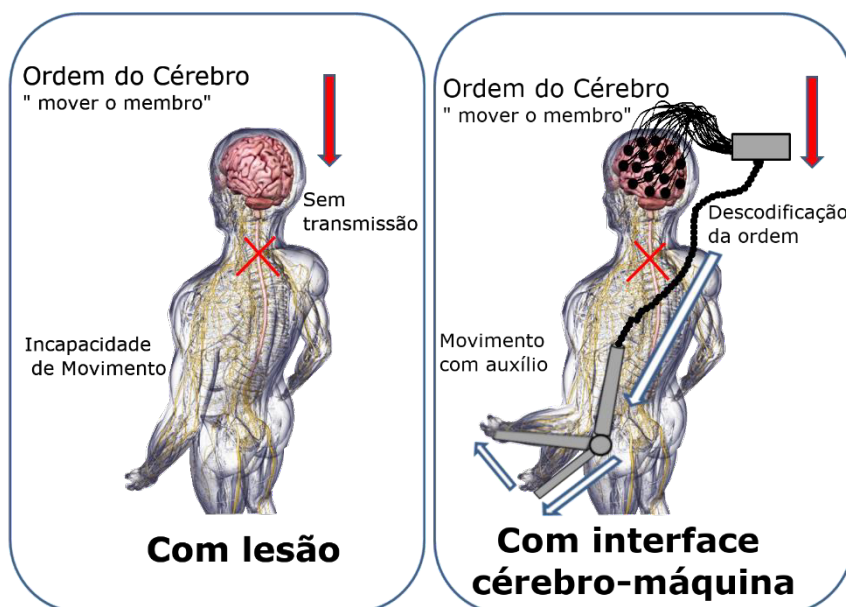


Figura 2 – Exemplo de paciente com lesão vertebro-medular antes e durante a utilização de uma interface cérebro-máquina.

As primeiras demonstrações empíricas do conceito de interface cérebro-máquina surgiram nas décadas de 60 e 70 do século XX (Dahl, 1962; Finley et al., 1975; Kamiya, 1971; Kaplan, 1975; Pelletier and Peper, 1977). No entanto, só no final desse mesmo século XX, com o acréscimo dos avanços tecnológicos registados na área da engenharia de materiais e nas ciências da computação, surge um enorme progresso no desenvolvimento das interfaces cérebro-máquina. Estes desenvolvimentos permitiram assim estudar, em tempo real e de forma extremamente detalhada, os circuitos neuronais associados à realização de múltiplas tarefas (Chapin et al., 1999; Carmena et al., 2003).

Foi através desta capacidade de descodificar em tempo real a atividade neuronal, que surgiu o desenvolvimento massivo das interfaces cérebro-máquina. Estas novas interfaces cérebro-máquina, substancialmente mais complexas do que aquelas que anteriormente haviam sido desenvolvidas nas décadas de 60 e 70 do século XX, permitiram a generalização da sua utilização em aplicações quotidianas e clínicas que começam agora a ser comercializadas. A possibilidade de interagir e interferir com a atividade cerebral não é nova, nem tão pouco, é específica das interfaces cérebro-máquina. No entanto, esta nova geração de engenhos trouxe a possibilidade de modular o sistema nervoso numa escala temporal e espacial bastante mais precisa do que aquela que se havia verificado anteriormente. Esta precisão teve um impacto tal, que se globalizou, prevendo-se que o mercado mundial destas interfaces se situe nos 1,24€ mil milhões em 2020 (Allied Market Research 2017).

Apesar destes avanços mostrarem ser extremamente relevantes para o desenvolvimento de soluções técnicas dirigidas a problemas clínicos de mobilidade, de visão, de comunicação, entre outros; há ainda alguma dificuldade em determinar de que forma o funcionamento do sistema nervoso, como um todo, é afetado aquando da utilização destes engenhos. É assim de prever que, neste contexto de risco/benefício; onde o primeiro parece ser largamente inferior ao segundo, surgirão inevitavelmente situações com potencial para que terceiros acedam a dados pessoais ou possam impor a realização de determinados atos a um sujeito, sem o seu consentimento informado e/ou conhecimento.

Neste momento, torna-se necessário salientar que as interfaces cérebro-máquina estão a deixar de ser meros equipamentos de reabilitação neurológica para se tornarem em equipamentos de uso quotidiano (Lebedev e Nicoletti, 2017). Por exemplo, já existem múltiplos sistemas para uso doméstico e aplicações de *faça-você-mesmo* (*'Do it yourself'*) para a elaboração de interfaces cérebro-máquina.

Neste texto descrevemos alguns aspetos técnicos das interfaces cérebro-máquina que consideramos ser importantes para uma discussão adequada do que pode ou não pode vir a ser um problema de responsabilidade para os múltiplos sujeitos envolvidos (destacando-se aqui os papéis do utilizador, do fabricante, e do pessoal clínico). Estes princípios são aqui descritos, com o objetivo de facilitar o trabalho do cientista do direito aquando do desenvolvimento do enquadramento legal a ser aplicado a esta tecnologia. As questões de natureza legal poderão depois ser abordadas pelos peritos da área em local apropriado tendo por base estas noções fundamentais.

Antes de explicarmos detalhadamente o que é uma interface cérebro-máquina explicaremos qual a sua relação com a corrupção. Depois descreveremos o que é o sentido de agência. Em seguida, descreveremos de que forma os componentes específicos das interfaces cérebro-máquina podem interferir com o sentido de agência. Ao longo desta descrição iremos salientar quais são os pontos mais críticos para o legislador ter em conta, assim como algumas medidas que devem ser tomadas para prevenir situações de corrupção aquando da utilização destes dispositivos.

2 – RELAÇÃO ENTRE INTERFACES CÉREBRO-MÁQUINA E CORRUPÇÃO

A ideia de corrupção surge geralmente associada à obtenção de uma vantagem indevida. Não desenvolveremos aqui esta noção, considerando apenas que esta existe enquanto forma de obter indevidamente “*vantagem patrimonial ou não patrimonial, ou a sua promessa, como contrapartida de acto ou de omissão contrários aos deveres*” de um determinado cargo, tal como descrita no Código Penal (Decreto-Lei nº 48/95 de 15-03-1995, Código Penal)².

O potencial de obtenção de vantagem indevida e as particularidades das interfaces cérebro-máquina são aqui abordados devido à presença de duas propriedades únicas destas interfaces. Primeiro, a utilização de dados provenientes do cérebro de um indivíduo permite o fácil acesso a informação privilegiada da qual o utilizador poderá não estar ciente de que estará a fornecer. Segundo, a manipulação de processos que ocorrem no cérebro humano tem o potencial de forçar um sujeito a realizar ações sem que tenha conhecimento e/ou consentimento das mesmas. No seu conjunto, estas duas situações são altamente potenciadoras de situações de vantagem indevida e de abusos de poder por parte de várias entidades.

² Não abordaremos aqui os detalhes da definição de corrupção por forma a podermos centrar o texto nas questões mais relacionadas com os detalhes da tecnologia em causa.

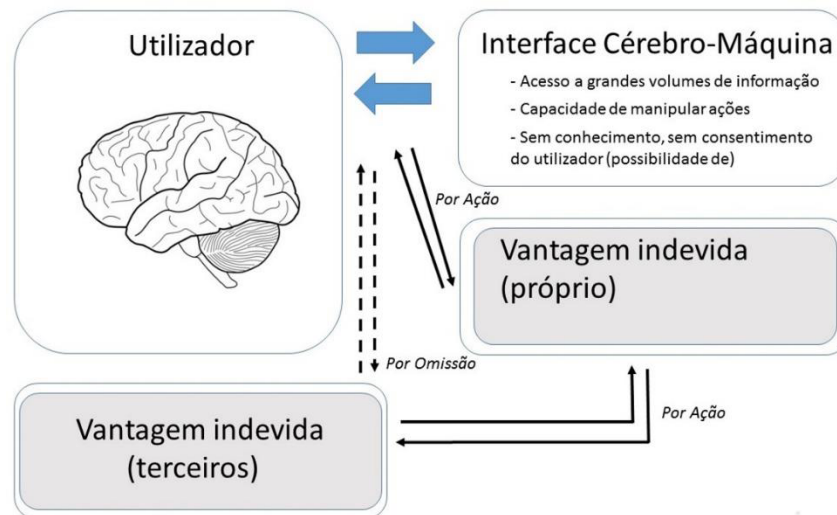


Figura 3 – Relação entre corrupção e interfaces cérebro-máquina. Como a manipulação das interfaces cérebro-máquina pode influenciar as ações do utilizador e/ou ter acesso a grandes volumes de informação, existe um enorme potencial para gerar situações de corrupção quer por ação quer por omissão.

Na Figura 3 apresentamos o esquema geral que relaciona o utilizador de uma interface cérebro-máquina com as situações passíveis de ocorrer corrupção. Um sujeito que tenha acesso à interface cérebro-máquina pode obter vantagem indevida por manipular a interface (obtendo vantagem para o próprio ou para terceiros), ou poderá, por omissão, permitir que terceiros obtenham essa mesma vantagem. Nos pontos seguintes demonstraremos em maior detalhe que aspetos técnicos devem ser considerados para que estas situações sejam prevenidas.

3 – DEFINIÇÃO DE INTERFACES CÉREBRO-MÁQUINA

Uma interface cérebro-máquina é um engenho que utiliza a atividade do sistema nervoso para controlar um determinado dispositivo ou função (Figura 4). Para poder permitir o controlo, recuperação ou expansão de uma função associada ao sistema nervoso, uma interface cérebro-máquina necessita, em geral, de dois elementos fundamentais: 1) uma forma de registar e analisar a atividade do sistema nervoso e 2) uma forma de concretizar a função desejada.³

³ É de salientar que no presente texto serão referidos trabalhos realizados tanto em humanos como em animais. Dada a rápida evolução que as interfaces cérebro-máquina têm apresentado; o espaço temporal que dista entre o desenvolvimento de uma interface para animais e a sua aplicação em

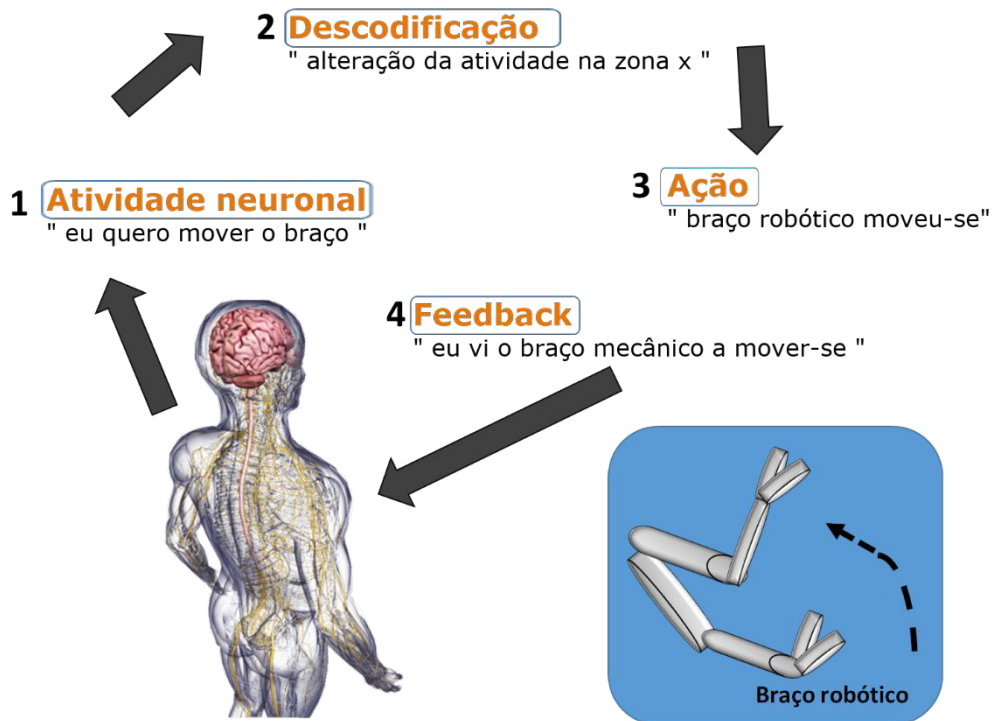


Figura 4 – Exemplo dos processos envolvidos numa interface cérebro-máquina. No exemplo apresentado, a interface cérebro-máquina regista atividade neuronal de um sujeito quando tem a intenção de ‘mover o braço’. Essa informação é analisada e descodificada em tempo real. A informação descodificada é utilizada para controlar uma máquina. O movimento realizado pela máquina (e observado pelo utilizador) vai servir de feedback (i.e. ‘o movimento foi o esperado’), permitindo controlar a qualidade do movimento que está a ser efetuado.

Veremos ainda que estes dois componentes gerais vão ser o resultado da interação das 8 variáveis específicas que condicionam a implementação das interfaces e que descreveremos em maior detalhe adiante. Se as interfaces cérebro-máquina precisam de interpretar e/ou modular a informação que está a ser processada no sistema nervoso, então quaisquer alterações que ocorram no sistema nervoso têm o potencial de afetar o funcionamento das interfaces cérebro-máquina. Mais precisamente, devemos ter como noção fundamental a considerar ao longo deste texto, que todas as

humanos tem sido de menos do que 5 anos. Como tal, indicaremos quais os trabalhos que foram realizados em humanos ou em animais, mas deveria estar sempre presente de que a transição entre as diferentes espécies poderá ser realizada num curto espaço de tempo.

variáveis que possam alterar a percepção que o indivíduo possui acerca da sua agência⁴, ou que tenham o potencial de alterar a sua tomada de decisão⁵ são relevantes para o estudo dos limites da responsabilidade aquando da utilização das interfaces cérebro-máquina.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (I):

- Uma vez que as interfaces cérebro-máquina têm o poder de alterar a atividade cerebral, também têm o poder de alterar a percepção que o indivíduo possui acerca da sua ação/agência e da sua responsabilidade.
- A possibilidade de manipulação das ações de um sujeito por parte de outrem sem que o sujeito esteja ciente dessa ocorrência potencia a ocorrência de situações de corrupção, uma vez que gera uma situação potencial de vantagem indevida ou abuso de poder sobre o utilizador.

PREVENÇÃO DA CORRUPÇÃO:

- Uma interface cérebro-máquina deverá ser sempre considerada à partida como um dispositivo com o potencial de alterar o sistema nervoso do utilizador num sentido indesejado (seja por abuso de poder ou por despoletar inadvertidamente uma patologia).
- Uma interface cérebro-máquina deverá ser sempre considerada à partida como um dispositivo com o potencial de ser pirateado ('hacked') por terceiros a fim de se extrair dados do utilizador ou a fim de se enviesar o funcionamento do cérebro e ações do utilizador.
- A comercialização de toda e qualquer interface cérebro-máquina deverá ser regulamentada e sujeita a um processo de legalização e aprovação prévia. Esta aprovação deverá incluir pareceres de uma equipa clínica e uma equipa de ciências da computação acerca da segurança clínica (por exemplo, interações com fármacos ou outras terapias) e segurança informática dos materiais, processamento e meios envolvidos. Também deverá ser elaborada uma tabela com os resultados esperados, possíveis efeitos secundários, e uma calendarização de quando se

⁴ O sentido de agência, que será discutido em maior detalhe adiante, diz respeito à noção que o sujeito tem de ter realizado uma ação.

⁵ Não discutiremos aqui os detalhes neurobiológicos da tomada de decisão. Apenas serão descritas algumas características do sentido de agência. Salienta-se, no entanto, que o sentido de agência interage de forma estreita com os mecanismos de tomada de decisão.

deverá monitorizar potenciais efeitos de neuroplasticidade a curto, médio e longo prazo (isto é, alterações provocadas pela utilização do engenho).

- A utilização de uma interface cérebro-máquina deverá ter resultar da concordância entre o utilizador/responsável legal, uma autoridade de saúde, uma autoridade de ciências da computação.
- A aprovação e regulamentação de cada dispositivo deverão incluir uma análise detalhada do percurso que a atividade neuronal/informação segue desde que é colhida até que é utilizada. Toda a atividade neuronal que não seja utilizada para a função legalmente descrita para a interface cérebro-máquina deverá seja descartada (isto é, não pode ser guardada e analisada), exceto mediante autorização expressa das entidades reguladoras e do utilizador e/ou responsável legal.

3.1 – COMPONENTES GERAIS

O primeiro componente, a forma de registo da atividade neuronal, pode ser conseguido através de uma multiplicidade de técnicas. Por exemplo, pode-se utilizar o registo por eletroencefalograma, o registo de potenciais de ação, o registo de potenciais de campo, ou mesmo de sinais que, não sendo exatamente atividade neuronal parecem constituir boas aproximações à mesma (por exemplo, o sinal BOLD⁶ na ressonância magnética funcional). Apesar de não se discutir aqui os detalhes destas técnicas, é de referir que elas podem apenas captar uma fração da atividade do sistema nervoso. Assim, cada uma destas técnicas estará associada a limitações e condicionantes específicas que podem alterar o resultado final da performance da interface cérebro-máquina (Lebedev e Nicolelis, 2017). Por outras palavras, é importante considerar sempre que *as interfaces cérebro-máquina interpretam uma fração da informação que foi registada e processada no sistema nervoso e, como tal, a sua ação é limitada à partida pelas condicionantes próprias de cada técnica.*

O segundo componente, a função realizada, pode ser obtida através de um leque bastante mais vasto de possibilidades do que o primeiro componente. Isto sucede porque o número de funções que podem ser concretizadas por uma interface cérebro-máquina, em potencial, é pelo menos tão grande quanto o número de funções que

⁶ BOLD é o acrónimo de Contraste Dependente do Nível de Oxigénio no Sangue (Blood-oxygen-level dependent contrast).

podem ser atribuídas ao sistema nervoso⁷. Nas suas versões mais conhecidas, as interfaces cérebro-máquina têm sido utilizadas para uma grande variedade de funções tais como recuperação da função motora (Donnati et al., 2016; Shokur et al., 2016), tratamento de epilepsia (Pais-Vieira et al., 2016), doença de Parkinson (Santana et al., 2014), recuperação de memórias (Deadwyler et al., 2013) entre muitas outras. No entanto, as interfaces cérebro-máquina começam já a ser utilizadas para um leque bastante mais vasto de funções que incluem a expansão do sistema nervoso com vista ao desenvolvimento de capacidades de processamento não existentes na natureza. De entre estas possibilidades salientam-se por exemplo, a comunicação direta entre cérebros (Pais-Vieira et al., 2013; Grau et al., 2014), o processamento consciente de sinais infravermelhos (Thomson et al., 2013), a realização de cálculos computacionais através de redes de cérebros interligados (Wang e Jung, 2011; Poli et al., 2013; Pais-Vieira et al., 2013; Yuan et al., 2013; Poli et al., 2014; Cecotti et al., 2014; Rao et al., 2014; Stocco et al., 2015; Ramakrishnam et al., 2015, Pais-Vieira et al., 2015; Matran-Fernandez et al., 2017), a implantação de memórias diretamente no cérebro (Deadwiler 2013), entre várias outras possibilidades (ver Lebedev e Nicolelis, 2017 para revisão). Em suma, o segundo componente indica-nos que *uma interface cérebro-máquina pode substituir ou completar funções do corpo humano, mas também pode realizar funções não existentes no corpo humano*.

4 – O SENTIDO DE AGÊNCIA

As interfaces cérebro-máquina têm a capacidade de modular a atividade do sistema nervoso com vista, por exemplo, à recuperação de funções. No entanto, as interfaces cérebro-máquina têm também o potencial de afetar outras funções que, em condições normais, nos permitem funcionar adequadamente em sociedade. Uma destas funções é o sentido de agência e está associado à capacidade do próprio se reconhecer como a origem das suas ações. A compreensão do sentido de agência é assim de extrema relevância para uma análise dos problemas associados à corrupção, porque o sujeito que vive em sociedade é considerado responsável pelos seus atos (Grübler, 2011; Haselager, 2013; Limerick et al., 2014; Marchesotti et al., 2017; Evans et al., 2015). Como tal, a modulação do sentido de agência, isto é, da noção que o próprio sujeito tem da sua responsabilidade sobre uma determinada ação, é de extrema relevância para a prevenção dos problemas associados à corrupção.

⁷ Adiante veremos que este número é bastante maior.

O sentido de agência (*Sense of Agency*) pode ser formalmente descrito como uma experiência de controlo das próprias ações na interação com o ambiente (Wegner 2002; Haggard e Chambon 2013; Moore e Flecher 2012; Kuhn et al., 2013, Renes et al., 2014).⁸ O sentido de agência só recentemente tem vindo a ser estudado de uma forma sistemática. A descrição das bases neurológicas deste sentido indica a existência de, pelo menos, dois mecanismos principais: 1) preparação para a ação, e 2) o feedback sensorial das nossas próprias ações. A título de exemplo, consideremos estes dois mecanismos no indivíduo saudável. Primeiro teríamos uma situação onde o sujeito apresentaria o pensamento de “*eu pretendo mover o braço para pegar num copo*” (isto é, a preparação para a ação), seguindo-se depois a experiência de “*vi e senti o copo quando o agarrei e ergui*” (isto é, feedback sensorial).

O mecanismo de feedback sensorial que resulta das nossas ações é mensurável através de um julgamento explícito acerca da atribuição de agência a si ou aos outros (Farrer et al., 2008; Synofzik et al., 2008)⁹. Portanto, o sujeito é, em geral, capaz de atribuir a realização da ação a uma entidade distinguindo em que situações o próprio, ou os outros, constitui causa da ação.

Na Figura 5, demonstramos como as interfaces cérebro-máquina utilizam frequentemente estes dois mecanismos. O primeiro, associado à ‘intenção’ de realizar uma ação, vai servir como base para decodificação do sinal relevante. O segundo, o ‘feedback’ das ações realizadas, vai contribuir largamente para a melhor performance da interface cérebro-máquina. Como se verá adiante, as interfaces cérebro-máquina podem atingir o seu resultado final através de múltiplos mecanismos, quer de decodificação, quer de feedback (sendo que este último até pode ser realizado diretamente no cérebro ou no sistema nervoso do indivíduo). Estes aspetos serão discutidos em maior detalhe aquando da apresentação das variáveis que interferem com as interfaces cérebro-máquina).

⁸ É importante referir aqui que a definição de sentido de agência não é a mesma para todos os autores. No entanto, enquanto base de estudo para o cientista do direito ou o legislador, a definição atual permite abarcar todas as variáveis de maior relevo.

⁹ É de salientar aqui, que alguns autores consideram que esta medida como tendo menos validade para avaliar mecanismos associados ao sentido de agência (fenómeno de ‘*temporal binding*’).

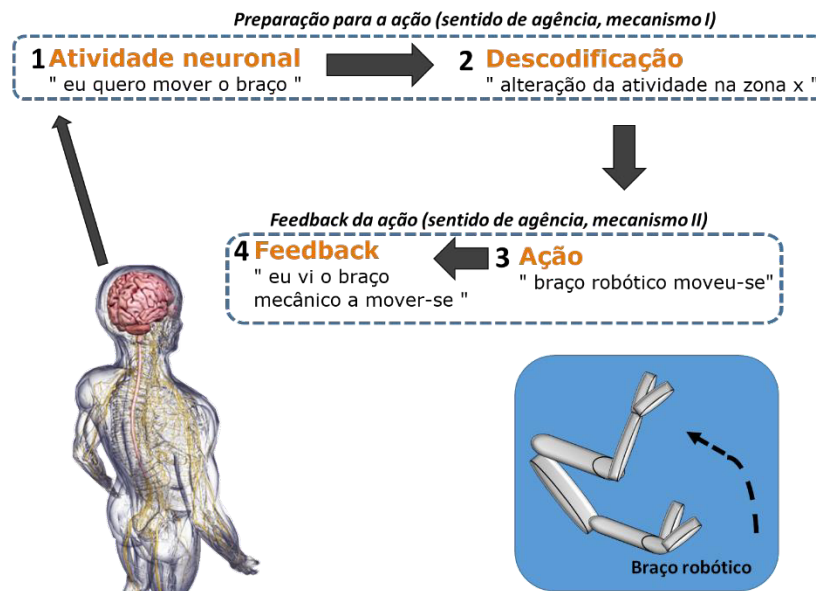


Figura 5 – O sentido de agência. Há dois mecanismos principais associados ao sentido de agência. O primeiro mecanismo diz respeito à preparação para a ação. O segundo mecanismo diz respeito ao feedback dessa mesma ação. As interfaces cérebro-máquina utilizam frequentemente estes dois mecanismos.

A neurobiologia do sentido de agência está associada a múltiplas regiões cerebrais. O conjunto destas regiões cerebrais e a sua conectividade (isto é, as múltiplas ligações estabelecidas entre elas)¹⁰ (Chambon et al., 2013; Haggard, 2017) têm sido repetidamente associadas a diferentes processos cognitivos e computacionais inerentes à experiência do sentido de agência.

As áreas que mais comumente surgem associadas ao sentido da agência em trabalhos experimentais são: o lobo frontal (córtex pré-motor, área motora suplementar; pré-frontal medial e lateral), o giro angular, parte do córtex parietal posterior e a ínsula (Figura 6). A ínsula, em particular a parte anterior, parece ter um papel preponderante surgindo de forma consistente associada ao relato de aumento da *'consciência'* (isto é, *awareness*) da ação causada pelo próprio (Farrer e Frith, 2002; ver também: Tsakiris et al., 2007; Farrer et al., 2003; Leube et al., 2003; Nahab et al., 2011; Karnath et al., 2005; Klein et al., 2007). Também relevantes para o presente texto são, o córtex parietal e o

¹⁰ A conectividade das áreas diz respeito à forma como elas podem ou não transferir informação entre elas. Veremos adiante que a conectividade está associada ao que descreveremos como a *arquitetura da rede*, e que tem um papel preponderante no processamento da informação.

lobo frontal, córtex medial, lateral e pré-motor que têm vindo a estar associados não só ao sentido de agência, mas também a situações onde o indivíduo que realizou a ação, julga que não o fez (isto é, atribui a ação a uma causa externa) (do inglês 'non-agency') (Sperduti et al., 2011; Farrer e Frith, 2002; Farrer et al., 2003; Chambon et al., 2013; Eimer & Schlaghecken, 2003; Nahab et al., 2017).

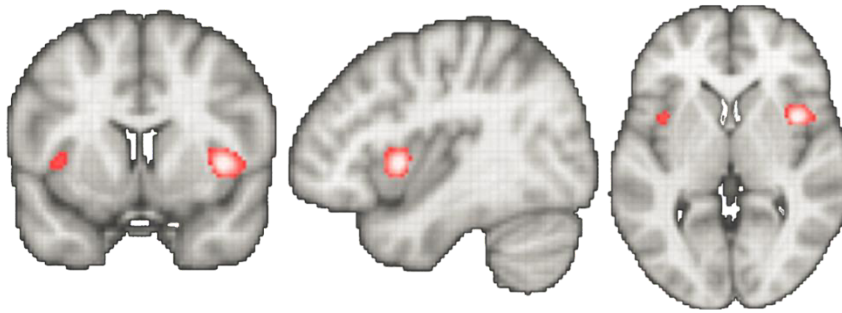


Figura 6 – A ínsula e o sentido de agência. Das múltiplas áreas envolvidas no processamento do sentido de agência a ínsula parece ser uma das mais relevantes. A figura ilustra as coordenadas $x=40$, $y=8$, $z=2$ do córtex insular anterior bilateral, ativada para atribuição de uma ação a si próprio (Farrer e Frith, 2002).¹¹ A região da ínsula está associada a muitas outras redes de processamento tais como a dor e a gustação. A manipulação da ínsula e de áreas associadas ao sentido de agência pode levar o sujeito a não ter noção de que realizou uma determinada ação (imagem proveniente de dados de CPV).

A descrição detalhada destas áreas é relevante porque indica que as interfaces cérebro-máquina que modulem as mesmas, têm o potencial de interferir com o sentido de agência. Isto significa que a modulação destas zonas pode fazer com que um sujeito que realizou uma determinada ação não tenha noção de que o fez. Vemos assim, que o sentido de agência parece ser constituído por, pelo menos, dois mecanismos principais, que têm por base múltiplas regiões cerebrais participantes em várias redes neuronais. Estas redes, por sua vez, interagem com múltiplas outras redes associadas direta ou indiretamente a estas funções. Além disso, verificámos que há um grande paralelismo entre o sentido de agência e os dois principais componentes das interfaces cérebro-

¹¹ Outras coordenadas do córtex insular associado a diferentes facetas do processamento do sentido de agência ($x = 40$, $y = -10$, $z = 16$; Farrer et al., 2003; 44, -10, 14 Tsakiris et al., 2007; $x = -32$, $y = 20$, $z = 2$; $x = -34$, $y = 16$, $z = 12$; $x = 21$, $y = 20$, $z = 5$; Nahab et al., 2011).

máquina (a técnica através da qual se colheu a atividade neuronal e a forma como a ação foi realizada).¹²

Relativamente ao sentido de agência, podemos assim concluir, que as interfaces cérebro-máquina que interferem direta ou indiretamente com o sentido de agência têm um enorme potencial de ser utilizadas para fins de obtenção de vantagem indevida (Haggard, 2017). Em seguida, descrevermos com um pouco de mais detalhe a constituição dos vários componentes de uma interface cérebro-máquina, salientando a sua relação com o sentido de agência e outras funções neurobiológicas.

5 – COMPONENTES ESPECÍFICOS DAS INTERFACES CÉREBRO-MÁQUINA QUE PODEM INTERFERIR COM A RESPONSABILIDADE

Mencionámos anteriormente que as interfaces cérebro-máquina são essencialmente constituídas por: 1) registo da atividade do sistema nervoso e, 2) a realização de uma ou mais funções.

Veremos agora que estes dois conjuntos principais resultam da interação entre oito variáveis de cariz mais técnico, mas de extrema relevância. São elas: 1) caracterização do nodo e número de nodos, 2) arquitetura da rede computacional, 3) características da técnica de registo, de descodificação, e de modulação; 4) natureza da tarefa a realizar; 5) tipo de resposta obtido, 6) grau de conhecimento dos nodos participantes, 7) volição e, 8) características da recompensa.

A descrição técnica aprofundada destas variáveis e das suas características neurobiológicas será realizada em local próprio, sendo realizada aqui apenas uma breve descrição de cada uma destas variáveis, por forma a identificar os aspetos relevantes para a prevenção da corrupção.

5.1 – NODO E NÚMERO DE NODOS

Começemos por assumir que toda e qualquer função realizada pelo sistema nervoso implica algum grau de processamento¹³. Se tal suceder, então podemos considerar que

¹² Veremos adiante que o feedback da ação realizada pode ser bastante mais complexo do que aquele que foi apresentado na figura. Por exemplo, pode-se gerar o feedback diretamente no cérebro do sujeito, sem qualquer necessidade de recorrer aos órgãos sensoriais.

¹³ Note-se que esta afirmação não implica que o sistema nervoso se resume a um computador, apenas indica que a atividade neuronal, enquanto base de uma determinada função pode ser associada, dentro de determinados limites, a algum tipo de função computacional. Não implicando essa propriedade que seja possível explicar a totalidade do funcionamento do sistema nervoso com base na metáfora computacional.

a interação entre o sujeito e qualquer interface cérebro-máquina poderá ser descrita como uma rede neuronal artificial. A descrição que se segue de rede neuronal exemplifica o propósito das redes neurais (Figura 7). Definiremos uma rede neuronal como um conjunto de nodos com ligações entre eles. O nodo seria então a unidade de computação de uma rede neuronal, e as ligações formadas entre eles seriam a direção e percurso que a informação percorre entre os vários nodos. Geralmente considera-se que há nodos de entrada (*input*), nodos de processamento e nodos de saída (*output*)¹⁴.

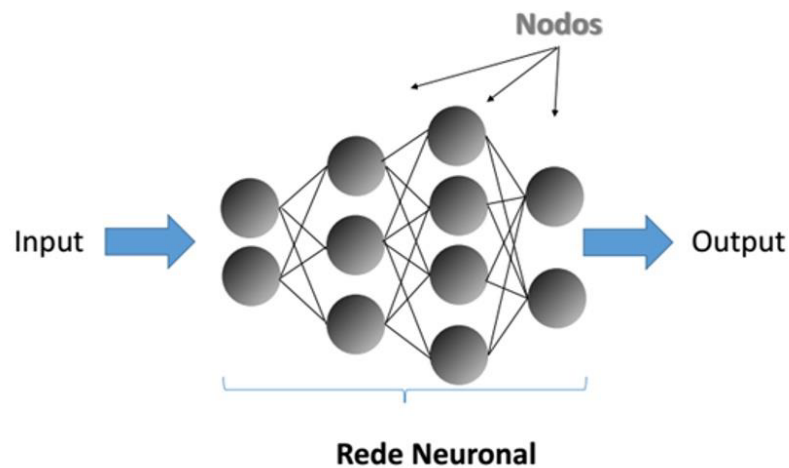


Figura 7 – Rede neuronal com *input*, nodos e *output*. Exemplo de uma rede neuronal constituída por múltiplos nodos. O *input* corresponde à entrada da informação, os nodos vão realizar o processamento (operações de computação) dessa mesma informação e, por último o *output* corresponderá à saída de informação.

A noção de nodo é aqui de extrema relevância porque, dependendo da interface cérebro-máquina em questão, pode assumir múltiplas formas. Por um lado, podemos ter uma rede neuronal integrada na interface cérebro-máquina. Por outro, podemos considerar que o conjunto do sujeito (ou sujeitos) que está a interagir com a interface constitui parte da própria rede neuronal. O que isto significa é que a ação da interface cérebro-máquina pode resultar da atividade neuronal de um ou mais indivíduos, de uma ou mais células, e de uma ou mais regiões cerebrais.

¹⁴ É de salientar que a arquitetura de redes pode assumir múltiplas outras formas bastante mais complexas do que aquelas que aqui são descritas.

Consideremos uma rede onde um único cérebro humano realiza a computação que permitirá a utilização da interface cérebro-máquina. Neste caso, podemos considerar que o cérebro humano será um nodo da rede. Consideremos agora um caso onde a rede utiliza a atividade média de dois cérebros, sem qualquer distinção entre as funções dos mesmos. Diríamos também que a nossa interface operava com base numa rede neuronal constituída por um único nodo (a unidade funcional), mas que esta resultava da junção da atividade de dois cérebros. Note-se que exemplos deste tipo de rede neuronal já foram testados anteriormente em humanos (Wang e Jung., 2011). No entanto, numa outra variante, poderíamos analisar a atividade destes dois cérebros em separado e depois calcular a média ponderada¹⁵ da sua atividade. Neste caso diríamos que a rede seria constituída por dois nodos, sendo que cada um deles corresponderia a um cérebro. É de salientar que, nos exemplos referidos, estivemos sempre a considerar o cérebro humano, mas que múltiplas outras variações podem ser utilizadas. Por exemplo, em alguns trabalhos anteriores a unidade de computação utilizada podia envolver em simultâneo a atividade de cérebros humanos e não-humanos, conjuntos de células, estruturas específicas, etc. (Li and Zhang, 2017, Koo et al., 2017; Yu et al., 2017; Carmena et al., 2003; Fuentes et al., 2009; Santana et al., 2014).

Conclui-se daqui, que é importante caracterizar os nodos porque estes determinam as condições de base que condicionam as decisões e ações tomadas, permitindo assim atribuir responsabilidades específicas a cada agente.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (II):

- Um dos componentes essenciais das interfaces cérebro-máquina é o nodo de processamento. O nodo de processamento constitui a unidade de computação e pode assumir várias formas e tipologias.

PREVENÇÃO DA CORRUPÇÃO:

- Aquando da análise dos algoritmos de processamento envolvidos na ação do interface cérebro-máquina, é necessária a caracterização dos nodos, porque esta

¹⁵ A média ponderada é uma média onde o peso que cada elemento tem no cálculo final da média pode ser diferente. Neste caso particular, a média ponderada permitiria que, quando a performance de um cérebro fosse melhor do que a do outro, a atividade proveniente deste cérebro tivesse um maior peso na ação final da interface cérebro-máquina.

caracterização determina as condições de base em que as decisões e ações são tomadas, permitindo uma clara definição de responsabilidades.

- Esta caracterização detalhada do papel e computação efetuada por cada um dos nodos evita que os dados colhidos durante a execução de tarefas possam ser utilizados para fins outros que aqueles que estão pré-estabelecidos.
- A prevenção da corrupção através da má utilização da computação realizada pelos nodos pode ser conseguida através da inventariação e da limitação da computação que pode ou não ser efetuada pelos nodos da rede.
- Sugere-se também a criação de uma classificação de algoritmos e interfaces cérebro-máquina. Estes grupos seriam formados de acordo com os pontos de fragilidade relativamente a possíveis intrusões ('hacking') ou manipulação por terceiros.

5.2 – ARQUITETURA DA REDE NEURONAL

Referimos anteriormente que, quando se utilizava um conjunto de cérebros ou por exemplo um conjunto de neurónios, ou regiões dentro do mesmo cérebro para realizar uma função; a forma como a informação era processada pela interface cérebro-máquina podia atribuir a cada nodo um peso e função específicos. A forma como os nodos se organizam e a forma como a informação flui entre os diferentes nodos denomina-se de *arquitetura da rede*. Inúmeros estudos na área da inteligência artificial têm descrito em detalhe a forma como a organização e ligações entre os nodos (assim como as funções de transferência associadas a cada uma dessas ligações) condicionam o processamento da informação (Fausset, 1994).

Um exemplo da forma como essas ligações podem ou não afetar a performance da interface cérebro-máquina é através de retro-propagação (*backpropagation*). Começamos por considerar duas redes. Numa delas os nodos processam a informação e não são informados do resultado final das suas ações. Na outra os nodos são informados acerca do resultado das suas ações (Pais-Vieira et al., 2013; Rao et al., 2014; Grau et al., 2014). Como facilmente se poderá perceber, os nodos que são informados das suas ações podem utilizar a informação resultante das suas ações para melhorar a sua performance.¹⁶

¹⁶ Veremos adiante que há ainda várias outras variáveis críticas dentro desta. De entre elas salienta-se o facto de o nodo estar ou não ciente de que está a praticar uma determinada ação.

Pode-se ainda referir aqui a situação onde há processamento paralelo (Figura 8 painel à direita) e processamento sequencial (Figura 8 painel à esquerda) (Pais-Vieira et al., 2015; Stocco et al., 2015). No processamento paralelo, os resultados de um determinado nodo (nodo A) não afetam os resultados do outro nodo (nodo B). No processamento sequencial o processamento de um nodo (nodo B) é feito com base nos resultados obtidos após o processamento de um outro nodo (nodo A) e são portanto dependentes.

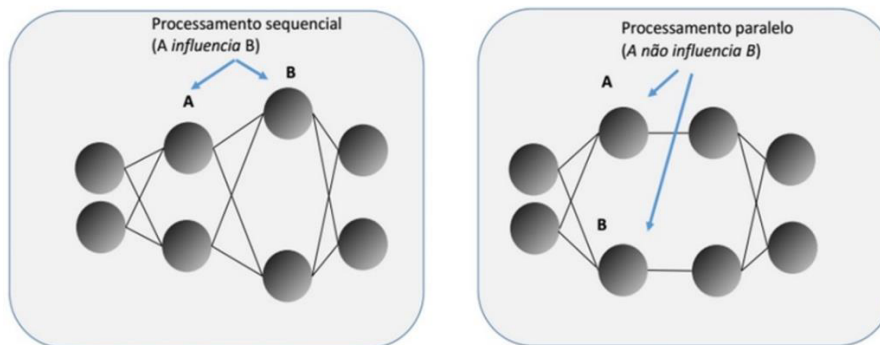


Figura 8 – Processamento sequencial e paralelo. No processamento sequencial a informação de um nodo (A) influencia o nodo seguinte (B), enquanto no processamento paralelo este processamento é independente (B é independente de A).

Em termos de estudo da responsabilidade individual, pode-se pensar que os sujeitos que atuam de forma independente terão responsabilidade apenas sobre as suas ações (e não sobre os restantes nodos)¹⁷. Já no caso do processamento em série, a má decisão por parte de um nodo na fase inicial será condicionadora de todas as decisões tomadas pelos nodos que se lhe seguem.

Por último, seria importante considerar que a intrusão de um terceiro, com vista à alteração do processamento da rede em cada um destes casos (processamento sequencial ou paralelo) também terá diferentes consequências.

¹⁷ Assumindo que os próprios estão cientes das suas ações.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (III):

- A ação realizada por um indivíduo que está a utilizar uma interface cérebro-máquina pode ser o resultado de um processamento realizado pelo próprio indivíduo, por outro indivíduo, ou por múltiplos indivíduos. Quando há múltiplos indivíduos envolvidos diz-se que a computação é feita em “rede”.
- A distribuição conhecida ou desconhecida, sequencial ou paralela de responsabilidades dentro de uma rede, constitui uma situação que pode potenciar a corrupção.

PREVENÇÃO DA CORRUPÇÃO:

- Os sujeitos que participam na utilização em rede de um interface cérebro-máquina necessitam ser informados das diferentes condições em que podem estar a processar informação.
- Quando necessário, esta informação pode apenas prever limites máximos e mínimos de participação de cada sujeito, sendo que a cada momento o sujeito poderá não ter conhecimento explícito da sua situação.
- O processamento em paralelo, dada a sua independência, pode servir como forma de controlo para alguns tipos de intrusão (‘hacking’) por terceiros.

5.3 – CARACTERÍSTICAS DA TÉCNICA DE REGISTO, DE DESCODIFICAÇÃO, E DE MODULAÇÃO

Como referido anteriormente, há um grande número de técnicas de registo e de estimulação do sistema nervoso que apresentam características e condicionantes diferentes (ver Lebedev e Nicolelis, 2017 para revisão). A precisão temporal e espacial de cada uma delas, assim como a sua invasividade¹⁸ são condicionantes fundamentais da sensibilidade da interface cérebro-máquina. O método de descodificação da atividade associada ao sistema nervoso será também um condicionante relevante para o funcionamento da interface cérebro-máquina. O problema geral associado aos algoritmos de descodificação prende-se com a capacidade que o algoritmo tem de discriminar entre várias classes de elementos.

¹⁸ Considera-se que uma técnica é invasiva quando requer a introdução de instrumentos ou substâncias no corpo (por exemplo, colocar uma prótese da anca ou injetar um fármaco). Considera-se que uma técnica é não-invasiva quando não requer a introdução de instrumentos ou substâncias no corpo (por exemplo, realizar a auscultação cardíaca com um estetoscópio, ou registar a atividade cerebral recorrendo a eletroencefalografia).

A título de exemplo, consideremos um algoritmo cujo objetivo seria a discriminação de sujeitos homens *versus* mulheres. A classe *Homem* apresentaria um conjunto de valores associados às variáveis, enquanto a classe *Mulher* apresentaria outros. Por exemplo, na Figura 9 apresentamos os valores de homens (*laranja*) e mulheres (*azul*) para as variáveis peso e altura. Apesar de os grupos serem globalmente diferentes, cada caso particular apresentaria necessariamente características de uma e de outra classe em maior ou menor grau. Assim, de um conjunto de dados onde se estudaria o peso, altura, índice de massa muscular, etc.; o algoritmo de discriminação deveria ser capaz de distinguir entre casos de homens e mulheres apresentados aleatoriamente. Um bom algoritmo deveria ter flexibilidade suficiente para poder classificar um homem de cabelo comprido e com baixo índice de massa muscular como pertencendo à classe *Homem* ou uma mulher de cabelo curto e com elevado índice de massa muscular como pertencendo à classe *Mulher* (apesar de se saber que estas características estão tendencialmente mais associadas ao grupo oposto). Na Figura 9 podemos ver que, apesar de as classes serem diferentes, há situações em que (unicamente com base nas características estudadas) a classificação destes dois casos específicos poderia ser difícil para o algoritmo de discriminação.

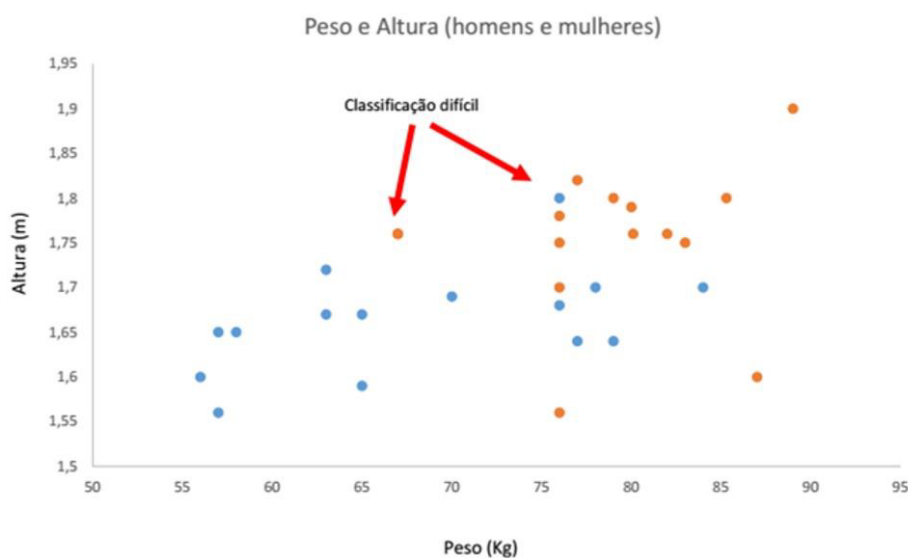


Figura 9 – Discriminação de classes por algoritmos. Apesar de os casos de homens ('laranja') e mulheres ('azul') apresentados na figura terem, em média, valores de altura e peso diferentes há bastante sobreposição. A classificação, por parte de um algoritmo, pode-se tornar particularmente complicada quando o caso em questão apresenta valores sugestivos de uma outra classe. As setas vermelhas indicam dois destes casos.

A importância de se ter em conta o algoritmo de classificação, aquando do estudo da responsabilidade, prende-se com estas situações limite onde o algoritmo implementado poderá não ser sempre capaz de identificar corretamente os elementos pertencentes a cada um dos grupos.

No exemplo anterior referimos a discriminação entre o género do sujeito. No entanto, poderíamos considerar agora uma situação onde o algoritmo estaria a tentar discriminar entre a atividade cerebral que promove ou inibe uma determinada ação. Se o sujeito que utiliza a interface cérebro-máquina estiver sob stress intenso, ou se tiver ocorrido algum tipo de alteração temporária na atividade cerebral, o algoritmo pode deixar de ser capaz de efetuar a discriminação entre os dois estados relevantes. Devemos então reter a noção de que os algoritmos de classificação apresentam um determinado grau de falibilidade. É importante determinar *a priori* em que medida *o grau de falibilidade de cada algoritmo pode levar a ações indesejáveis/inesperadas durante o uso da interface cérebro-máquina*. Este grau de falibilidade é particularmente relevante por um lado, no que diz respeito à segurança do utilizador (i.e. 'prevenção de ações inadequadas'), e por outro porque constitui uma situação suscetível à intrusão por terceiros (i.e. de incerteza de resultados).

Além das técnicas de registo e de descodificação é importante considerar também as técnicas de modulação do sistema nervoso.¹⁹ Estas técnicas são relevantes porque, em alguns casos, a interface cérebro-máquina não só regista a atividade do cérebro, como também vai induzir diretamente alterações na mesma²⁰. Uma situação onde a modulação da atividade neuronal é extremamente importante é no controlo de sintomas. Múltiplas interfaces têm sido desenvolvidas com vista à redução de, por exemplo, convulsões na epilepsia, e de tremores na doença de Parkinson. No entanto, da mesma forma que a modulação (seja por aumento ou redução da atividade neuronal num conjunto de áreas) pode ser utilizada para fins de melhoria clínica, também poderá ser utilizada com outros fins. Por exemplo, já foi anteriormente demonstrada em animais e humanos a possibilidade de geração de movimentos automáticos num sujeito apenas através do controlo da atividade cerebral que teria como origem a atividade cerebral de um outro sujeito (i.e. interfaces cérebro-cérebro) (Pais-Vieira et al., 2013;

¹⁹ Entende-se por *modulação* como qualquer alteração (por exemplo, aumento ou redução) na atividade de uma determinada zona. As interfaces cérebro-máquina podem analisar as modulações que ocorrem na atividade cerebral, mas também podem induzir modulações na atividade cerebral.

²⁰ Recorde-se que as interfaces cérebro-máquina podem recolher informação do sistema nervoso, mas também podem introduzir alterações no seu funcionamento.

Shanechi et al., 2014; Grau et al., 2014; Rao et al 2014). Estes estudos indicam que a modulação do sistema nervoso pela interface cérebro-máquina pode ser sujeita a intrusão (*'hacking'*) por parte de terceiros. *Como tal, a interface cérebro-máquina tem o potencial para gerar atividade em regiões do cérebro, modificando as sensações, crenças, ações, pensamentos, vontades e decisões do sujeito com ou sem o seu conhecimento e com ou sem o seu consentimento.*

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (IV):

- Entre a intenção do sujeito, as modulações que constituem a atividade neuronal e a ação observada, existe tipicamente algum grau de incerteza associado ao algoritmo e à técnica de registo.
- A modulação da atividade neuronal tem o potencial para modificar as sensações, crenças, ações, pensamentos, vontades e decisões do sujeito com ou sem o seu conhecimento, e com ou sem o seu consentimento.
- A qualidade da função das interfaces cérebro-máquina é determinada, em grande parte, pela capacidade de as técnicas utilizadas permitirem uma adequada aquisição de sinal e modulação das regiões pretendidas.

PREVENÇÃO DA CORRUPÇÃO:

- Os parâmetros utilizados nos algoritmos a implementar nas interfaces cérebro-máquina deverão ser validados por uma entidade governamental, para que se possa garantir a segurança do utilizador sob um ponto de vista de proteção de dados pessoais e de manipulação não-autorizada dos processos cerebrais.
- Da mesma forma, dever-se-á proceder a uma avaliação cuidadosa da segurança em termos de potencial promoção de alterações neurológicas e neuropsiquiátricas.

5.4 – AÇÃO ITERATIVA OU CONTÍNUA

A tarefa a realizar, ou a função a desempenhar podem-se apresentar de múltiplas formas. No caso de uma interface cérebro-máquina que controla um membro robótico, a função de controlo será tipicamente contínua, isto é, o sujeito vai realizando um movimento (ou conjunto de movimentos) que perduram no tempo permitindo-lhe fazer correções quando necessário (Donnatti et al., 2016; Shokur et al., 2016). Já em outros casos, a interface cérebro-máquina poderá estar associada à entrega de um pacote de

informação (por exemplo, estimulação durante 30 minutos) que irá resultar numa alteração do sistema (Fuentes et al., 2009; Santana et al., 2014). Em termos de responsabilidade pelas ações realizadas, é importante reter aqui que, no primeiro caso, o conjunto de ações se prolonga no tempo, sendo possível ao sujeito corrigir a sua ação inicial. Já no segundo caso, da administração de um pacote único de informação, poder-se-ia em teoria, gerar uma alteração de estado no sistema, fazendo com que o indivíduo perdesse o controlo da agência ou tomada de decisão²¹. Por outras palavras, a avaliação da responsabilidade de uma ação por parte de um sujeito está associada, entre outras variáveis, ao tempo que dista entre a realização da mesma e os seus efeitos. Note-se que esta distância entre o estímulo e a ação constitui assim um fator de incerteza para o utilizador e, como tal, torna-se num ponto de fragilidade onde se poderá tentar adulterar o funcionamento normal do sistema.

Afirmámos anteriormente que, em teoria, as interfaces cérebro-máquina apresentam o potencial de realizar pelo menos o mesmo número de funções que o sistema nervoso humano. Na verdade, o número de funções a realizar é provavelmente bastante superior ao referido, uma vez que não é necessário que as interfaces cérebro-máquina realizem tarefas que estejam associadas ao sistema nervoso. Por exemplo, trabalhos anteriores já demonstraram redes de cérebros a realizar processos computacionais tais como previsão meteorológica (Pais-Vieira et al., 2015) ou a identificação de fontes de radiação infravermelha (Thomsom et al., 2013), que não correspondem às funções tipicamente associadas ao sistema nervoso. O que isto implica é que a computação formal efetuada pelo sistema nervoso pode ser utilizada de múltiplas formas. Referimos acima que a intrusão (*'hacking'*) por parte de um sujeito poderia fazer com que o utilizador da interface realizasse uma função que não desejasse. Aqui surge ainda uma outra possibilidade, a possibilidade de o utilizador estar convencido de que está a realizar uma determinada função e, na verdade estar a realizar outra. Consideremos, a título de exemplo, que o utilizador está totalmente convencido de que está a utilizar um braço robótico e que deve utilizar a sua atividade cerebral para o braço robótico se mover num determinado espaço (por exemplo, tocar num ponto no espaço virtual). Um intruso (*'hacker'*) poderia condicionar este paradigma experimental, de tal forma a que a tarefa realizada pelo utilizador correspondesse exatamente ao percurso que um engenho explosivo deveria seguir para atingir um determinado alvo. O utilizador estaria

²¹ Por exemplo, numa das primeiras demonstrações de interfaces cérebro-cérebro em humanos a ação motora do segundo sujeito para disparar um canhão virtual era automática e totalmente controlada pela atividade neuronal do primeiro sujeito, não podendo ser corrigida (Rao et al., 2014).

convencido de que a computação realizada por si correspondia ao exercício normal de aperfeiçoar a utilização da sua interface. No entanto, o intruso apropriar-se-ia do processamento cerebral realizado para obter a realização, em tempo real, de uma outra função por ele desejada. Este tipo de situação seria semelhante ao Project Pigeon (Orcon) liderado por Skinner onde pombos eram ensinados a controlar misseis (Skinner, 1959).

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (V):

- O intervalo de tempo que decorre entre uma determinada modulação e o comportamento daí resultante constitui uma variável crucial para a possível intrusão ('hacking') de terceiros.
- Quando a ação da interface cérebro-máquina é iterativa, ela decorre em blocos, não permitindo qualquer tipo de correção após o seu início. Quando a ação é contínua, pode ser possível corrigi-la após o seu início.
- A informação utilizada para a realização de uma função pode ser comum a outras funções formalmente semelhantes. Este tipo de situação é particularmente suscetível à intrusão ('hacking') por terceiros, uma vez que não é possível ao utilizador saber que a sua atividade cerebral está a ser utilizada para duas ou mais funções em simultâneo.

PREVENÇÃO DA CORRUPÇÃO:

- Não há uma estratégia única para a resolução da possível intrusão de terceiros aquando da decisão de quais os intervalos de tempo a utilizar em cada tipo de interface cérebro-máquina. Esta variável deverá ser analisada para cada caso individual.
- O problema da utilização indevida da computação realizada numa tarefa para a realização de uma tarefa formalmente semelhante reside essencialmente na segurança informática. Cabe ao fabricante, garantir que a interface cérebro-máquina não pode ser pirateada por terceiros durante a sua utilização.
- Quando a interface cérebro-máquina estiver associada a ações que possam ser críticas para a segurança do utilizador e da sociedade (por exemplo, levar a lesão ou dano grave do próprio ou outrem) o fabricante deverá utilizar sempre o algoritmo que evite a ação e/ou que permita a correção da mesma após o seu início.

5.5 – NATUREZA DA TAREFA A REALIZAR

A natureza da tarefa a realizar será um condicionante do resultado final, na medida em que formas diretas e indiretas de computação podem ser utilizadas. Começemos por considerar o caso em que a atividade neuronal associada ao movimento do membro superior é utilizada, com conhecimento de todos os participantes, para controlar o movimento do membro superior de um indivíduo com lesão (adaptado de Shانهchi et al., 2014). Neste caso a natureza da computação realizada tem aplicação direta no tipo de tarefa que se espera que a interface realize. Por outro lado, podemos pensar no caso onde a atividade neuronal resultante de uma determinada área cerebral associada ao processamento tátil de um indivíduo seria utilizada para ativar uma zona motora, com vista à realização de uma tarefa de percepção tátil (Pais-Vieira et al., 2015). Neste segundo caso poderíamos dizer que a estimulação efetuada (por exemplo, na zona motora) teria como objetivo a geração de uma resposta num outro local que poderia estar mais ou menos relacionado com a função em causa. Para o leitor não versado em neurociências podemos estabelecer um paralelismo com uma situação do nosso quotidiano. Consideremos que dois indivíduos estão a conversar. Na forma direta podemos pensar como sendo a situação em que um diz abertamente a palavra ‘*sim*’ em resposta a uma questão feita pelo outro. Já na forma indireta podemos pensar num acenar de cabeça que, por convenção, significa também “*sim*”. Nesta forma indireta, o resultado final é o mesmo, mas apenas houve uma diferença na forma como a informação foi processada. No caso das interfaces cérebro-máquina, podemos pensar na modulação de uma determinada zona do cérebro que, apesar de não ser preferencialmente dedicada a um tipo de informação permitirá atingir o resultado final esperado.

A importância desta classificação direta e indireta é que a segunda poderá estar mais sujeita ao aparecimento de enviesamentos (porque é mais incerta) e portanto mais sujeita à intrusão por terceiros sem que o utilizador se aperceba disso. Além disso, uma vez que a forma indireta exige mais processamento, poderá requerer um período de aprendizagem/adaptação mais longo (uma vez que a resposta obtida não é a que irá ser realizada na tarefa). Este aumento no tempo de processamento e de aprendizagem poderá ser condicionante do resultado final de duas formas: se for requerida uma resposta imediata, poderá ser atrasada; por outro lado, se for necessário algum tipo de deliberação, este tempo adicional poderá permitir tomar a decisão correta. Conclui-se assim, que cada uma destas situações poderá ser positiva ou negativa consoante o cenário em que se verifica.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (VI):

- A realização de funções específicas de uma determinada tarefa poderá ser afetada por manipulação cerebral ou utilização da informação colhida em áreas consideradas “não específicas” para essa mesma função.
- A utilização de atividade cerebral proveniente de regiões primárias ou associativas pode alterar o tempo de processamento.

PREVENÇÃO DA CORRUPÇÃO:

- A utilização de áreas não-específicas parece estar mais suscetível a atos de corrupção porque a modulação das mesmas vai implicar a modulação de outras áreas possivelmente não relacionadas com o objetivo esperado.
- Sempre que possível dever-se-á utilizar áreas primordialmente associadas à função que se deseja realizar. É de salientar, no entanto, que em muitas situações é mais fácil utilizar redes neuronais que envolvem múltiplas áreas do que utilizar informação (ou modular) apenas uma única área.
- Recomenda-se aqui uma análise detalhada de cada caso, por parte das autoridades clínica e informática, sendo que o utilizador deverá ficar a par dos riscos associados.

5.6 – ENVOLVIMENTO DO SUJEITO

A relação estabelecida entre o utilizador e a interface cérebro-máquina pode ocorrer de múltiplas formas. O grau de envolvimento do utilizador, isto é o estado de consciência que o utilizador tem da interface, constitui um ponto de fragilidade do sistema e, como tal condiciona situações de corrupção.

Para melhor explicar esta variável consideremos três exemplos onde o grau de envolvimento do utilizador é bastante diferente.

No primeiro caso, a forma mais simples, podemos considerar um sujeito que está a controlar um braço robótico através da sua atividade cerebral e que recebe feedback táctil virtual acerca desses movimentos. O utilizador treinado está totalmente ciente de: 1) qual o tipo de atividade neuronal que a interface está a registar (por exemplo, registar a atividade do córtex pré-motor), 2) qual o tipo de processamento realizado (por exemplo, tentar prever a intenção do movimento desejado para um braço robótico), e 3) qual o tipo e efeito de modulação realizada (por exemplo, receber feedback tátil como forma de corrigir os movimentos do braço robótico).

Podemos agora considerar um segundo caso, totalmente diferente, onde um sujeito utiliza uma interface cérebro-máquina para controlo de epilepsia. O controlo da epilepsia pode envolver múltiplas áreas, podendo assim considerar-se que esta interface foi inicialmente colocada com um objetivo específico, mas cuja forma de atuação pode ser mais complexa. Neste caso particular, pode ser necessário registar e/ou modular múltiplas áreas em simultâneo, a fim de que o efeito ocorra a nível da rede (que é composta por múltiplas regiões cerebrais) e não necessariamente numa única área. Nesta situação o sujeito terá, provavelmente, um conhecimento geral das áreas que estão a ser registadas e moduladas, e terá também um conhecimento parcial do tipo de processamento realizado. Provavelmente, terá ainda um conhecimento parcial das modulações entregues (i.e. alterações de sensação que lhe podem sugerir que algumas regiões foram moduladas, mas não de outras), e terá conhecimento dos objetivos gerais (i.e. reduzir o aparecimento de convulsões). No entanto, o sujeito provavelmente não terá conhecimento dos objetivos locais mais técnicos (tais como, aumentar num período curto a excitação de uma zona particular que se sabe ser inibitória de uma segunda zona, que por sua vez afeta o restante da rede em questão).

Num terceiro cenário, um pouco mais crítico, podemos mesmo considerar uma interface cérebro-máquina que esteja a ser utilizada num indivíduo que se encontra em coma induzido (por exemplo, para reduzir a atividade de uma determinada zona)²². Neste caso (independentemente da função realizada) o sujeito não terá conhecimento: das áreas registadas e moduladas, do processamento realizado, dos objetivos globais ou locais, nem sequer da sua participação.

Verificamos então, nestes três casos diferentes (controlo voluntário de um braço robótico, modulação de múltiplas áreas para controlo de epilepsia, e modulação de uma área num paciente em coma), que há múltiplos graus de conhecimento que o indivíduo pode ter acerca do seu papel no funcionamento da interface cérebro-máquina. Tal como demonstrado na tabela 1, a ausência ou não de conhecimento pode ser relativamente à área registada e/ou modulada, ao processamento realizado, à participação na utilização da interface, às modulações recebidas, e relativamente aos objetivos locais e globais.

²² Em determinadas situações clínicas é necessário manter o paciente em coma induzido para, por exemplo, se estabilizar determinadas funções.

Área registada	Processamento realizado	Área modulada	Participação	Modulações entregues	Objetivos locais e/ou globais
Conhecida	Conhecida	Conhecida	Conhecida	Conhecida	Conhecidos
Desconhecida	Desconhecida	Desconhecida	Desconhecida	Desconhecida	Desconhecidos
Parcialmente Conhecida	Parcialmente Conhecida	Parcialmente Conhecida	Parcialmente Conhecida	Parcialmente Conhecida	Parcialmente Conhecidos

Tabela 1: Grau de conhecimento do utilizador relativamente às variáveis intervenientes na utilização de uma interface cérebro máquina.

Apesar de estes três casos ilustrarem situações aparentemente simples, levantam questões relevantes quando se tem em conta o crescimento exponencial que se tem vindo a observar no mercado das interfaces cérebro-máquina. Por exemplo, com o aumento da possibilidade de intrusão (*'hacking'*) da atividade cerebral é necessário considerar a possibilidade da intrusão na modulação de grupos de cérebros sem que estes tenham total conhecimento de quais os procedimentos que estão a realizar.

Uma situação algo parecida (mas de forma alguma ilegal) e que pode ajudar a compreender este fenómeno é a resolução dos *Captcha*²³ (Figura 10). Quando se tenta aceder a um portal na internet, somos deparados frequentemente com um conjunto de letras e números com posições e formatos deturpados que devemos identificar. O objetivo é distinguir se a entidade que está a aceder ao portal é um humano ou um programa.



Figura 10 – Exemplo de CAPTCHA utilizado em portais da internet. A interpretação de múltiplas fontes, cores e orientações de algarismos e letras tende a constituir um desafio maior para as máquinas do que para os humanos.

Não é raro que os indivíduos que estão a utilizar um portal na internet (Bursztein et al., 2010) também estejam a servir como sujeitos para a realização de estudos sobre cognição. Isto é, o tempo que o indivíduo demora a responder, os erros que comete, e

²³ Captcha é acrónimo de *Completely Automated Public Turing test to tell Computers and Humans Apart*; que significa Teste de Turing Totalmente Autónomo para Distinguir Humanos de Computadores.

em alguns casos a forma como gasta mais tempo a olhar ou a manipular o cursor numa determinada área do ecrã, são bons indicadores da forma como o indivíduo está a processar a informação. Transpondo agora esta noção para as interfaces cérebro-máquina, podemos então pensar que surgirão situações onde o sujeito, ou um grupo de sujeitos, estarão a utilizar uma interface cérebro-máquina, com um objetivo aparentemente claro e do conhecimento de todos os participantes, mas onde na realidade há todo um conjunto de outras informações a serem colhidas e processadas. Denominaremos a esta situação de “*computação inocente*”.

Por exemplo, em insetos, roedores e primatas já se testou a utilização de redes computacionais de cérebros com vista à realização de tarefas (Ramakrishnam et al. 2015, Pais-Vieira et al., 2013; 2015; Grau et al., 2014; Stocco et al., 2015). De acordo com o conhecimento atual, não se prevê que os sujeitos tivessem conhecimento da maior parte das variáveis aqui descritas (área registada, processamento realizado, participação). Quando estas atividades envolviam humanos ou primatas a área analisada era desconhecida (uma vez que era registada a atividade de múltiplas áreas) e o processamento realizado era parcialmente conhecido (i.e. o resultado final tipicamente dependia de vários indivíduos e poderia envolver algum tipo de ponderação). Nestes casos, os objetivos globais eram geralmente conhecidos, uma vez que as tarefas envolviam geralmente algum tipo de recompensa.

Assim, verificamos que as interfaces cérebro-máquina colhem e analisam uma grande quantidade de informação, podendo inadvertidamente obter, também, informação relacionada com parâmetros que não foram autorizados pelo sujeito.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (VII):

- Uma interface cérebro-máquina colhe e analisa grandes volumes de informação, utilizando geralmente apenas uma pequena parte.
- Uma interface cérebro-máquina pode atuar no sistema nervoso do indivíduo sem que este tenha algum tipo de experiência indicativo dessa ação.
- Devido à natureza computacional do sistema nervoso, o sujeito que utiliza a interface cérebro-máquina pode processar dados sem que tenha qualquer noção de que está a realizar computação inocente.

PREVENÇÃO DA CORRUPÇÃO:

- Os algoritmos ou engenhos que potenciem a computação inocente deverão ser evitados sempre que possível.
- A fim de proteger o utilizador de realizar computação inocente, os algoritmos aprovados para a utilização de interfaces cérebro-máquina deverão descartar toda a informação que não estiver relacionada com a função para a qual a interface cérebro-máquina está descrita.
- As autoridades competentes deverão emitir um documento específico para cada classe de interfaces cérebro-máquina a indicar qual a informação que é recolhida, o tipo de modulação efetuada, quais as zonas registadas e/ou moduladas, bem como os efeitos esperados a curto, médio e longo prazo.
- Dever-se-ão indicar também possíveis sinais e sintomas que possam estar associados à intrusão por parte de terceiros (e.g. sintomas neurológicos fora do comum, alterações no tempo de processamento).

5.7 – RECOMPENSA

A utilização das interfaces cérebro-máquina está geralmente associada a um fim específico, tal como a recuperação de uma função afetada por patologia. O sucesso deste tipo de utilização geralmente está associado a algum tipo de recompensa. Por exemplo, para o sujeito que perdeu uma determinada função (e.g. perda de locomoção associada a paraplegia), a capacidade de voltar a realizar essa mesma função pode ser uma recompensa de elevado valor em si mesma (e.g. voltar a andar). Por outro lado, poderá haver outras formas de recompensa de índole bastante diferente. Pensemos, por exemplo, na capacidade de expandir o tipo ou número de funções que o corpo humano habitualmente realiza, tais como a aquisição de perceção do espectro infravermelho (Thomson et al., 2013) ou do campo geomagnético (Norimoto e Ikegaya, 2015) que, poderão constituir, por si só, uma forma de recompensa, uma vez que permitem o acesso a um vasto conjunto de novas experiências. A importância da variável recompensa aquando da utilização das interfaces cérebro-máquina é também relevante, porque pode condicionar a capacidade de aprendizagem do utilizador (Shokur et al., 2017, Donnatti et al., 2017) e, até certo ponto, a sua volição.

No entanto, é de extrema importância perceber que a recompensa pode surgir através da interpretação que o sujeito faz do facto de ter atingido um determinado objetivo, mas também pode surgir por modulação direta do sistema de recompensa.

Aquando da realização de uma tarefa conjunta, tal como se verifica para as redes de cérebros a funcionar em conjunto, pretende-se atingir um determinado objetivo. A realização deste objetivo coincidirá tipicamente com algum tipo de recompensa. Será de prever que quando se proceder à realização conjunta de atividades através da utilização de interfaces cérebro-máquina poderá suceder que a distribuição da recompensa pelos múltiplos intervenientes seja diferenciada (isto é, ser diferente para cada elemento), ou que pelo menos seja percecionada pelos elementos como tal. Este facto tem, por si só, o potencial para alterar (de forma consciente ou inconsciente) a avaliação que os sujeitos farão da sua responsabilidade e da sua motivação para a realização de uma função conjunta. Assim, *o nodo que considerar que a recompensa recebida é desajustada ao esforço realizado para a obtenção da mesma terá maior probabilidade de alterar a sua forma de participação na realização de uma tarefa conjunta.*

Um outro tipo de recompensa é aquele que é obtido por modulação direta do sistema de recompensa. Este tipo de modulação, à semelhança do que foi descrito para outras espécies (Olds e Milner, 1954), consiste na obtenção da recompensa através modulação da atividade cerebral sem necessidade de administração de qualquer substância ou sem interferência de qualquer ação externa. No limite, este tipo de interface cérebro-máquina poderá levar a adições graves com as consequentes alterações de comportamentos. Tal como em algumas doenças atuais (por exemplo, o abuso de substâncias), é possível que uma única experiência de recompensa seja suficiente para alterar de forma prolongada o comportamento do sujeito. Assim, torna-se relevante para o presente estudo, a noção de que *a modulação direta do sistema de recompensa através do uso de interfaces cérebro-máquina poderá levar a alterações graves das estruturas associadas aos circuitos que influenciam a responsabilidade.*

Por último, devemos levantar aqui uma última questão que poderá surgir aquando da manipulação da recompensa durante a utilização de interfaces cérebro-máquina. Uma vez que é possível que uma interface cérebro-máquina funcione sem qualquer conhecimento do seu utilizador (tal como vimos anteriormente), torna-se também possível que a recompensa atribuída ao utilizador também seja entregue sem o seu conhecimento. O que isto significa é que o utilizador pode estar a ser manipulado para realizar uma determinada função, sem ter conhecimento disso, mas como está a ser recompensado, pode apresentar apenas uma sensação geral de “bem-estar”. Este tipo de situação permitiria que um terceiro tivesse um controlo de funções do sujeito, pura e simplesmente pelo facto de que aquele conjunto de ações faria com que o sujeito se

“*sentisse bem*”. Um exemplo do nosso cotidiano que ajuda a explicar esta situação seria o ato de “*correr*” como forma de exercício. Começamos por assumir que alguns indivíduos têm prazer em correr e outros não. Poderíamos pensar no caso do sujeito que, não gosta de correr, mas que através de modulação de determinadas áreas passa a “*ter vontade*” de realizar essas ações. Quando realiza essas ações, ele “*sente-se bem*”, porque o seu sistema de recompensa está a ser modulado pela interface cérebro-máquina. Vemos que, em nenhum momento, o indivíduo tem noção de estar a ser manipulado no sentido de realizar determinadas ações ou para as experiências de prazer que delas decorrem. No entanto, todo este sistema está orientado para que o sujeito (de forma inconsciente) reforce este tipo de ações no futuro. Surge assim, como questão fundamental neste ponto, a noção de que a recompensa depende de múltiplos fatores e que pode ser atribuída sem que o sujeito se aperceba da mesma. Este tipo de situação é altamente potenciador de situações de abuso de poder e de manipulação indevida das ações do sujeito.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (VIII):

- Se o nodo perceber o esforço realizado como sendo maior do que a recompensa obtida, haverá maior probabilidade de este alterar a sua atividade neuronal no futuro, no sentido de uma pior performance.
- A modulação do sistema de recompensa envolve direta ou indiretamente as mesmas estruturas que participam na responsabilidade que o sujeito tem das suas ações.
- É possível manipular o comportamento dos sujeitos e recompensá-los por essas mesmas ações sem que os sujeitos alguma vez tenham conhecimento pleno de que foram manipulados ou recompensados.

PREVENÇÃO DA CORRUPÇÃO:

- A modulação de regiões cerebrais associadas ao sistema de recompensa tem um enorme potencial de criar novos tipos de patologia e/ou de serem utilizadas para fins criminosos, uma vez que constituem um potente móbil para a ação e responsabilidade do sujeito.
- O sujeito que utiliza interfaces cérebro-máquina capazes de modular o seu sistema de recompensa deverá estar ciente da existência dessas modulações ou, no caso de tal não ser possível (por exemplo, coma), dever-se-á obter a autorização do

responsável legal (por exemplo, família) em conjunto com uma opinião da equipa clínica.

- A introdução de dispositivos que interfiram diretamente com o sistema de recompensa constitui um dos pontos críticos a serem considerados na utilização de interfaces cérebro-máquina dado o elevado potencial para o aparecimento de patologias e de abuso de poder por parte de terceiros. Como tal, a modulação deste sistema deverá ser sempre requerer uma monitorização especialmente cuidadosa.

5.8 – VOLIÇÃO

A volição, de acordo com Haggard (2008), consiste no conjunto de decisões que guiam a decisão e o tempo da ação. Tal como indicado no início deste texto, todas as variáveis que possam alterar a perceção que o indivíduo possui acerca da sua agência, ou que tenham o potencial de alterar a sua tomada de decisão serão relevantes para o estudo da responsabilidade associada à utilização das interfaces cérebro-máquina.

O papel da volição está bastante associado aos circuitos da recompensa. As interfaces cérebro-máquina podem interferir com a volição de forma direta ou indireta. Tal como referido anteriormente, a interferência direta estaria relacionada com a modulação de áreas que se sabe controlarem diretamente a volição, enquanto a interferência indireta estaria associada à modulação de zonas que teriam a possibilidade de afetar as zonas associadas à volição (Janicak e Dokucu, 2015). Tal como referido para as outras variáveis, esta interferência pode, ou não, ser percebida pelo sujeito e pode, ou não, interferir com as restantes variáveis.

INFORMAÇÃO RELEVANTE PARA O LEGISLADOR (IX):

- Uma interface cérebro-máquina pode interferir com a volição do sujeito de forma direta ou indireta.

PREVENÇÃO DA CORRUPÇÃO:

- À semelhança dos efeitos da recompensa, o processamento cerebral associado à volição constitui um ponto crítico que deverá requerer a autorização do indivíduo e/ou família. Este tipo de modulação deverá ser sujeito à avaliação cuidadosa por uma equipa clínica.

- A intrusão de terceiros aquando da utilização de interfaces cérebro-máquina capazes de modular zonas associadas à volição é particularmente suscetível a abusos de poder, porque alteram os mecanismos básicos que condicionam a ação do indivíduo em sociedade.

6 – NOTAS FINAIS

O presente texto abordou as principais variáveis que estão associadas ao funcionamento de uma interface cérebro-máquina sendo referidos pontos críticos onde situações de corrupção e abuso de poder podem ocorrer. São também propostas aqui um conjunto de medidas que se espera que venham a aumentar a segurança do utilizador e a transparência nas decisões tomadas. Também se prevê que estas medidas limitem o poder e responsabilizem cada um dos intervenientes que têm o potencial de enviesar as ações e funções cerebrais do sujeito.

Como forma de prevenir o conjunto de situações descritas ao longo do texto, sumariamos aqui as medidas gerais a ter em conta aquando da elaboração do enquadramento legal destas tecnologias.

NOÇÕES GERAIS

Uma interface cérebro-máquina é um dispositivo com o potencial de alterar o sistema nervoso do utilizador por abuso de poder ou por potenciação de patologia. Este dispositivo pode também ser pirateado podendo ser extraídos dados do utilizador sem o seu consentimento, ou levando a que o sujeito realize determinadas ações (ainda que o mesmo não tenha conhecimento delas).

ENTIDADES RESPONSÁVEIS

A comercialização das interfaces cérebro-máquina deverá ser regulamentada, sendo que a inclusão de pareceres de uma equipa clínica e uma equipa de ciências da computação deverá ser sempre incluída. Este processo deverá assim incluir sempre três entidades em concordância: o utilizador/responsável legal, uma autoridade de saúde, uma autoridade de ciências da computação. Uma vez que há uma grande quantidade de técnicas de registo e estimulação neuronal, e uma enorme variedade de algoritmos a aplicar a aprovação de cada dispositivo deverá incluir uma análise do processamento da informação, sendo que toda a atividade neuronal que não participar da função para a qual o equipamento está aprovado deverá ser descartada.

Aquando da análise dos algoritmos de processamento envolvidos na ação da interface cérebro-máquina, é necessária a caracterização dos nodos (i.e. das unidades de processamento). Será importante criar uma classificação de algoritmos e interfaces cérebro-máquina em grupos que se sabe apresentarem mais ou menos pontos de fragilidade relativamente a possíveis intrusões (*'hacking'*) ou manipulação por terceiros.

Relativamente à entidade clínica dever-se-á sempre fazer uma avaliação da segurança atual e futura relativamente ao aparecimento de alterações neuroplásticas e outras resultantes da utilização devida e indevida da interface cérebro-máquina.

UTILIZADORES EM REDE

Os sujeitos que participam na utilização em rede de uma interface cérebro-máquina necessitam ser informados das diferentes condições em que podem estar a processar informação. Quando necessário, esta informação pode apenas prever limites máximos e mínimos de participação de cada sujeito, sendo que a cada momento o sujeito poderá não ter conhecimento explícito da sua situação.

INTRUSÃO ('HACKING')

O ónus da segurança informática de cada interface cérebro-máquina caberá essencialmente ao fabricante, sendo que a autoridade em ciências da computação apenas deverá confirmar que o dispositivo foi desenvolvido de acordo com as melhores práticas de segurança disponíveis à data. Prevê-se, no entanto, que este seja o ponto de de mais difícil resolução no futuro.

UTILIZAÇÃO DE ÁREAS NÃO-ESPECÍFICAS

A utilização de áreas não-específicas parece estar mais suscetível a atos de corrupção porque a modulação das mesmas vai implicar a modulação de outras áreas possivelmente não relacionadas com o objetivo esperado. Isto faz com que o meio através do qual se atingiu um fim, se torne menos claro.

Sempre que possível dever-se-á utilizar áreas primordialmente associadas à função que se deseja realizar. É de salientar, no entanto, que em muitas situações é mais fácil utilizar redes neuronais que envolvem múltiplas áreas do que utilizar informação (ou modular) apenas uma única área. A autoridade clínica em conjunto com a autoridade informática deverá pronunciar-se sobre os riscos associados à utilização de áreas não-específicas para a função em questão, uma vez que estas parecem estar mais suscetíveis a atos de corrupção e a alterações neurológicas imprevisíveis.

SISTEMA DE RECOMPENSA E VOLIÇÃO

A modulação de regiões cerebrais associadas ao sistema de recompensa ou à volição tem um enorme potencial de criar novos tipos de patologia e/ou de serem utilizadas para fins criminosos. A introdução de dispositivos que interfiram diretamente com o sistema de recompensa ou a volição constitui um dos pontos críticos a serem considerados na utilização das interfaces cérebro-máquina e, como tal, deverá ser sempre cuidadosamente monitorizada.

USO NÃO-REGULAMENTADO

Os autores preveem que o uso não-regulamentado das interfaces cérebro-máquina, nomeadamente através de situações de *faça-você-mesmo*, atinja valores elevados pouco tempo após a generalização das interfaces cérebro-máquina. Esta situação será, provavelmente, de difícil controlo e poderá levar a consequências graves. Propõem-se aqui a realização de campanhas de sensibilização junto do grande público a fim de promover a não utilização de tais dispositivos.

7 – CONCLUSÕES

A massificação da utilização das interfaces cérebro-máquina tem o potencial para levar a múltiplas situações de corrupção que, pela complexidade dos mecanismos e técnicas envolvidos podem não ser de fácil resolução. A prevenção dessas situações deve combinar esforços de diferentes entidades da área da saúde e das ciências da computação, no sentido de analisar em detalhe cada uma das principais variáveis que condicionam a utilização das interfaces cérebro-máquina.

REFERÊNCIAS

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, 98(20), 11818-11823.

Bursztein, E., Bethard, S., Fabry, C., Mitchell, J.C. and Jurafsky, D., 2010, May. How good are humans at solving CAPTCHAs? A large scale evaluation. In *Security and Privacy (SP), 2010 IEEE Symposium on* (pp. 399-413). IEEE.

Carmena, J.M., Lebedev, M.A., Crist, R.E., O'Doherty, J.E., Santucci, D.M., Dimitrov, D.F., Patil, P.G., Henriquez, C.S. and Nicolelis, M.A., 2003. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS biology*, 1(2), p.e42.

Cecotti H, Rivet B. Subject combination and electrode selection in cooperative brain-computer interface based on event related potentials. *Brain Sci.* 2014 Apr 30;4(2):335-55. doi: 10.3390/brainsci4020335.

Chambon V., Wenke D., Fleming S. M., Prinz W., Haggard P. (2013). An online neural substrate for sense of agency. *Cereb. Cortex* 23, 1031–1037

Chapin, J.K., Moxon, K.A., Markowitz, R.S. and Nicolelis, M.A., 1999. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature neuroscience*, 2(7), pp.664-670.

Dahl, W.D., 1962. An alpha rhythm feedback control unit (No. NRL-5848). Naval Research Laboratory Washington D.C.

Deadwyler SA, Berger TW, Sweatt AJ, Song D, Chan RH, Opris I, Gerhardt GA, Marmarelis VZ, Hampson RE. Donor/recipient enhancement of memory in rat hippocampus. *Front Syst Neurosci.* 2013 Dec 26;7:120.

Donati AR, Shokur S, Morya E, Campos DS, Moioli RC, Gitti CM, Augusto PB, Tripodi S, Pires CG, Pereira GA, Brasil FL, Gallo S, Lin AA, Takigami AK, Aratanha MA, Joshi S, Bleuler H, Cheng G, Rudolph A, Nicolelis MA. Long-Term Training with a Brain-Machine Interface-Based Gait Protocol Induces Partial Neurological Recovery in Paraplegic Patients. *Sci Rep.* 2016 Aug 11;6:30383.

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological psychology*, 64(1-2), 7-26.

Evans, N., Gale, S., Schurger, A., & Blanke, O. (2015). Visual feedback dominates the sense of agency for brain-machine actions. *PloS one*, 10(6).

Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M (2003). Modulating the experience of agency: a positron emission tomography study. *NeuroImage*, 18:324–33.

Farrer, C., Frey, S.H., Van Horn, J.D., Tunik, E., Turk, D., Inati, S., and Grafton, S.T. (2008). The angular gyrus computes action awareness representations. *Cerebr. Cortex* 18, 254–261.

Farrer, C., Frith, C.D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* 15 (3), 596–603

Fausett, L.V., (1994). *Fundamentals of neural networks: architectures, algorithms, and applications* (Vol. 3). Englewood Cliffs: prentice-Hall.

Finley, W.W., Smith, H.A. and Etherton, M.D., 1975. Reduction of seizures and normalization of the EEG in a severe epileptic following sensorimotor biofeedback training: Preliminary study. *Biological psychology*, 2(3), pp.189-203.

Fuentes, R., Petersson, P., Siesser, W.B., Caron, M.G. and Nicoletis, M.A., 2009. Spinal cord stimulation restores locomotion in animal models of Parkinson's disease. *Science*, 323(5921), pp.1578-1582.

Grübler, G. (2011). Beyond the responsibility gap. Discussion note on responsibility and liability in the use of brain-computer interfaces. *AI & society*, 26(4), 377.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196.

Haggard, P., & Chambon, V. (2013). Sense of agency. *Current Biology*, 22(10), R390-R392.

Haggard, P., 2008. Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), p.934.

Haselager, P. (2013). Did I do that? Brain–computer interfacing and the sense of agency. *Minds and Machines*, 23(3), 405-418.

Janicak, P.G. and Dokucu, M.E., 2015. Transcranial magnetic stimulation for the treatment of major depression. *Neuropsychiatric disease and treatment*, 11, p.1549.

Kamiya, J., 1971. Biofeedback training in voluntary control of EEG alpha rhythms. *California medicine*, 115(3), p.44.

Kaplan, B.J., 1975. Biofeedback in epileptics: Equivocal relationship of reinforced EEG frequency to seizure reduction. *Epilepsia*, 16(3), pp.477-485.

Karnath HO, Baier B, Nagele T. 2005. Awareness of the functioning of one's own limbs mediated by the insular cortex? *J Neurosci*. 25:7134—7138

Klein TA, Endrass T, Kathmann N, Neumann J, von Cramon DY, Ullsperger M. (2007). Neural correlates of error awareness. *Neuroimage*. 34:1774---1781

Koo B, Koh CS, Park HY, Lee HG, Chang JW, Choi S, Shin HC. Manipulation of Rat Movement via Nigrostriatal Stimulation Controlled by Human Visually Evoked Potentials. *Sci Rep*. 2017 May 24;7(1):2340.

Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, 49(7), 1935-1942.

Lebedev MA, Nicolelis MA. Brain-Machine Interfaces: From Basic Science to Neuroprostheses and Neurorehabilitation. *Physiol Rev*. 2017 Apr;97(2):767-837

Leube DT, Knoblich G, Erb M, Grodd W, Bartels M, Kircher TT (2003). The neural correlates of perceiving one's own movements. *Neuroimage*, 20:2084–90.

Li G, Zhang D. Brain-Computer Interface Controlled Cyborg: Establishing a Functional Information Transfer Pathway from Human Brain to Cockroach Brain. *PLoS One*. 2016 Mar 16;11(3):e0150667.

Limerick, H., Coyle, D., & Moore, J. W. (2014). The experience of agency in human-computer interactions: a review. *Frontiers in human neuroscience*, 8, 643.

Marchesotti, S., Martuzzi, R., Schurger, A., Blefari, M. L., del Millán, J. R., Bleuler, H., & Blanke, O. (2017). Cortical and subcortical mechanisms of brain-machine interfaces. *Human brain mapping*, 38(6), 2971-2989.

Matran-Fernandez, A. and Poli, R., 2017. Towards the automated localisation of targets in rapid image-sifting by collaborative brain-computer interfaces. *PLoS One*, 12(5), p.e0178498.

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Consciousness and cognition*, 21(1), 59-68

Nahab FB, Kundu P, Gallea C, Kakareka J, Pursley R, Pohida T, et al. (2011). The neural processes underlying self-agency. *Cereb Cortex*, 21:48–55.

Nahab, F.B., Kundu, P., Maurer, C., Shen, Q. and Hallett, M., 2017. Impaired sense of agency in functional movement disorders: An fMRI study. *PloS one*, 12(4), p.e0172502.

Norimoto, H., & Ikegaya, Y. (2015). Visual cortical prosthesis with a geomagnetic compass restores spatial navigation in blind rats. *Current Biology*, 25(8), 1091-1095.

Olds, J. and Milner, P., 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology*, 47(6), p.419.

Pais-Vieira M, Lebedev M, Kunicki C, Wang J, Nicolelis MA. A brain-to-brain interface for real-time sharing of sensorimotor information. *Sci Rep*. 2013;3:1319.

Pais-Vieira M, Chiuffa G, Lebedev M, Yadav A, Nicolelis MA. Building an organic computing device with multiple interconnected brains. *Sci Rep*. 2015 Jul 9;5:11869.

Pais-Vieira M, Yadav AP, Moreira D, Guggenmos D, Santos A, Lebedev M, Nicolelis MA. A Closed Loop Brain-machine Interface for Epilepsy Control Using Dorsal Column Electrical Stimulation. *Sci Rep*. 2016 Sep 8;6:32814.

Pelletier, K.R. and Peper, E., 1977. Developing a biofeedback model: Alpha EEG feedback as a means for pain control. *International Journal of Clinical and Experimental Hypnosis*, 25(4), pp.361-371.

Poli, R., Cinel, C., Matran-Fernandez, A., Sepulveda, F. and Stoica, A., 2013, March. Towards cooperative brain-computer interfaces for space navigation. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (pp. 149-160).

Poli, R., Valeriani, D. and Cinel, C., 2014. Collaborative brain-computer interface for aiding decision-making. *PloS one*, 9(7), p.e102693.

Ramakrishnan A, Ifft PJ, Pais-Vieira M, Byun YW, Zhuang KZ, Lebedev MA, Nicolelis MA. Computing Arm Movements with a Monkey Brainet. *Sci Rep*. 2015 Jul 9;5:10767

Rao RP, Stocco A, Bryan M, Sarma D, Youngquist TM, Wu J, Prat CS. A direct brain-to-brain interface in humans. *PLoS One*. 2014 Nov 5;9(11):e111332.

Renes, R. A., van Haren, N. E., Aarts, H., & Vink, M. (2014). An exploratory fMRI study into inferences of self-agency. *Social cognitive and affective neuroscience*, 10(5), 708-712.

Santana, M.B., Halje, P., Simplício, H., Richter, U., Freire, M.A.M., Petersson, P., Fuentes, R. and Nicoletis, M.A., 2014. Spinal cord stimulation alleviates motor deficits in a primate model of Parkinson disease. *Neuron*, 84(4), pp.716-722.

Schwarz, D.A., Lebedev, M.A., Hanson, T.L., Dimitrov, D.F., Lehew, G., Meloy, J., Rajangam, S., Subramanian, V., Ifft, P.J., Li, Z. and Ramakrishnan, A., 2014. Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nature methods*, 11(6), p.670.

Shokur S, Gallo S, Moioli RC, Donati AR, Morya E, Bleuler H, Nicoletis MA. Assimilation of virtual legs and perception of floor texture by complete paraplegic patients receiving artificial tactile feedback. *Sci Rep*. 2016 Sep 19;6:32293.

Skinner, B.F., 1960. Pigeons in a pelican. *American Psychologist*, 15(1), p.28.

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self-and external-agency attribution: a brief review and meta-analysis. *Brain Structure and Function*, 216(2), 151-157.

Stocco A, Prat CS, Losey DM, Cronin JA, Wu J, Abernethy JA, Rao RP. Playing 20 Questions with the Mind: Collaborative Problem Solving by Humans Using a Brain-to-Brain Interface. *PLoS One*. 2015 Sep 23;10(9):e0137303

Synofzik, M., Vosgerau, G., Newen, A., 2008. Beyond the comparator model: a multifactorial two-step account of agency. *Conscious. Cogn.* 17 (1), 219–239.

Thomson, E.E., Carra, R. and Nicoletis, M.A., 2013. Perceiving invisible light through a somatosensory cortical prosthesis. *Nature communications*, 4, p.1482.

Tsakiris A, Hesse MD, Boy C, Haggard P, Fink GR (2007) Neural signatures of body ownership: a sensory network for bodily selfconsciousness. *Cereb Cortex* 17:2235–2244.

Wang Y, Jung TP. A collaborative brain-computer interface for improving human performance. *PLoS One*. 2011;6(5):e20422. doi: 10.1371/journal.pone.0020422.

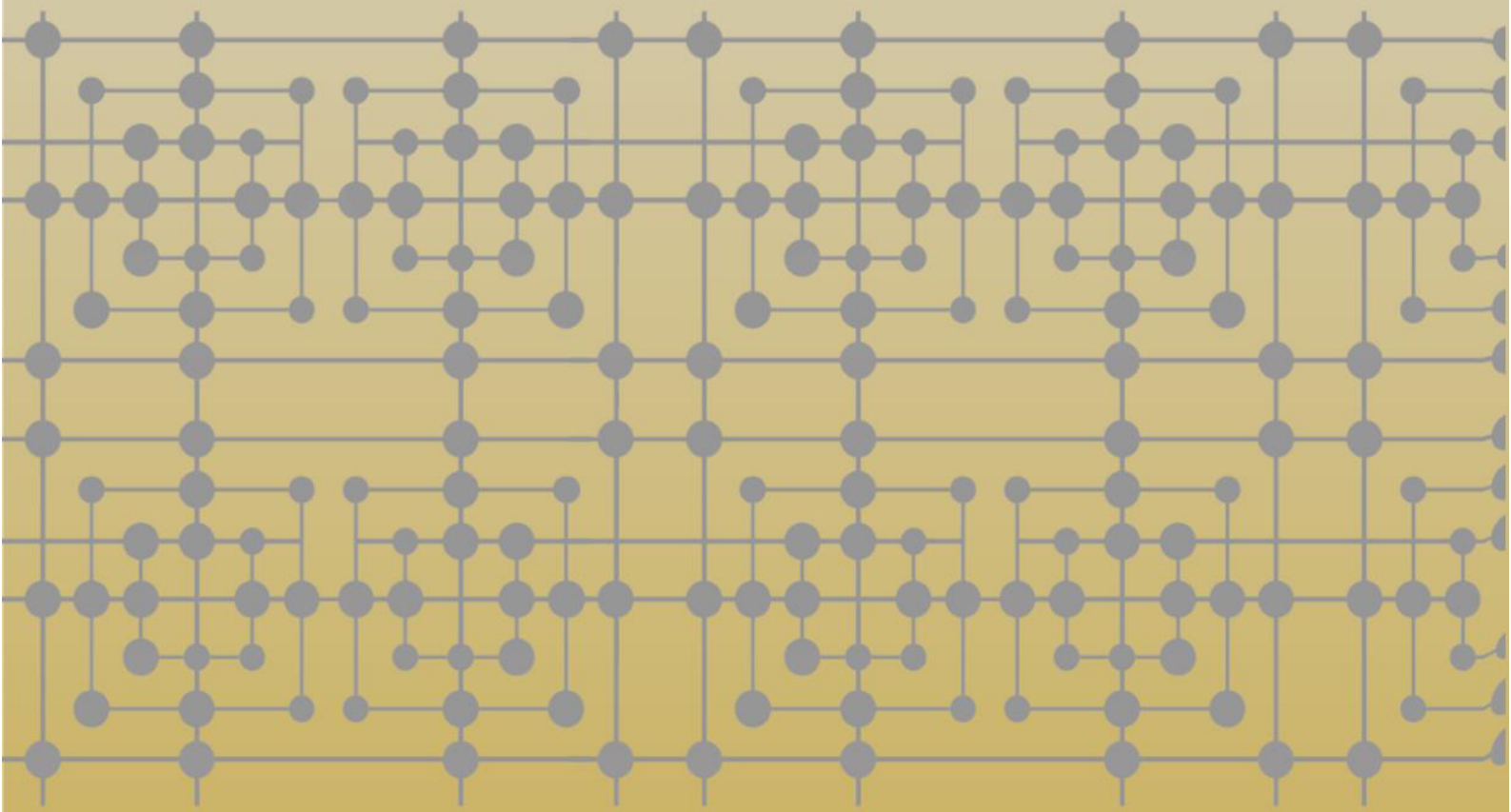
Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge: MIT Press.

Yu Y, Qian C, Wu Z, Pan G. Mind-controlled Ratbot: A Brain-to-brain System. 2014 IEEE International Conference on Pervasive Computing and Communications

Yuan, P., Wang, Y., Gao, X., Jung, T.P. and Gao, S., 2013, July. A collaborative brain-computer interface for accelerating human decision making. In Proceedings of the 7th international conference on Universal Access in Human-Computer Interaction: design methods, tools, and interaction techniques for eInclusion-Volume Part I (pp. 672-681). Springer-Verlag.

Portais consultados:

- 1 Allied Market Research 2017, <https://www.alliedmarketresearch.com/brain-computer-interfaces-market> (consultado em 12/12/2017)
- 2 Decreto-Lei nº 48/95 de 15-03-1995, Artigo 372.º Corrupção passiva para acto ilícito, http://bdjur.almedina.net/item.php?field=item_id&value=80262 (consultado em 30/05/2018)



CONSELHO DE
PREVENÇÃO DA
CORRUPÇÃO