

UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO

CAMPUS MARÍLIA

FACULDADE DE FILOSOFIA E CIÊNCIAS (FFC)

LUIZ AUGUSTO ROSA

A CONSCIÊNCIA SEGUNDO DAMÁSIO:

UM ENFOQUE NEUROFILOSÓFICO

MARÍLIA

2018

UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO

CAMPUS DE MARÍLIA

FACULDADE DE FILOSOFIA E CIÊNCIAS (FFC)

PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

LUIZ AUGUSTO ROSA

A CONCIÊNCIA SEGUNDO DAMÁSIO:

UM ENFOQUE NEUROFILOSÓFICO

Dissertação de Mestrado apresentada no programa de Pós-graduação em Filosofia da Faculdade de Filosofia e Ciências da Universidade Estadual Júlio de Mesquita Filho – UNESP – Campus de Marília para a obtenção do título de Mestre em Filosofia. Linha: Ciência Cognitiva, Filosofia da Mente e Semiótica.

Orientador: Prof. Dr. Jonas Gonçalves Coelho

MARÍLIA

2018

Rosa, Luiz Augusto.
R788c A consciência segundo Damásio: um enfoque
neurofilosófico / Luiz Augusto Rosa. – Marília, 2018.
138 f. ; 30 cm.

Orientador: Jonas Gonçalves Coelho.
Dissertação (Mestrado em Filosofia) – Universidade
Estadual Paulista (Unesp), Faculdade de Filosofia e
Ciências, 2018.
Bibliografia: f. 130-138

1. Damásio, António R. – 1944- . 2. Consciência. 3.
Emoções. 4. Self (Filosofia). 5. Neurobiologia. 6. Filosofia
da mente. I. Título.

CDD 128.2

Ficha catalográfica elaborada por
André Sávio Craveiro Bueno
CRB 8/8211
Unesp – Faculdade de Filosofia e Ciências

LUIZ AUGUSTO ROSA

**A CONSCIÊNCIA SEGUNDO DAMÁSIO:
UM ENFOQUE NEUROFILOSÓFICO**

Dissertação de Mestrado para a obtenção do título de Mestre em Filosofia

BANCA EXAMINADORA:

Prof. Dr. Jonas Gonçalves Coelho

Prof. Dr. Alfredo Pereira Junior

Prof. Dr. Osvaldo Frota Pessoa Junior

Marília, 02 de fevereiro de 2018

AGRADECIMENTOS

Eu agradeço:

Ao Prof. Dr. Jonas Gonçalves Coelho pela sua constante orientação e apoio, fundamental tanto para os resultados preliminares e finais do trabalho quanto para minha formação acadêmica como um todo.

Aos Professores Doutores Alfredo Pereira Junior e Osvaldo Frota Pessoa Junior pelos conselhos, sugestões e críticas construtivas dadas a versões preliminares do trabalho, as quais foram fundamentais para a construção dessa versão final do trabalho.

Ao parecerista anônimo da FAPESP, responsável por avaliar versões preliminares e finais do trabalho, dos quais as sugestões também foram fundamentais para a construção dessa versão final do trabalho.

Ao meu irmão, Luan A. R. Falchette, por seu constante apoio e suporte.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo fundamental apoio financeiro dado a esse projeto.

Resumo

Nessa dissertação buscamos fazer uma análise filosófica das teorias sobre a emoção, sentimento, *self* e consciência de António Damásio frente ao contexto neurobiológico contemporâneo. No primeiro capítulo, nós fizemos uma revisão de literatura sobre os dois paradigmas da consciência vigentes na neurobiologia, no caso, o paradigma cognitivista, onde revisamos os trabalhos de Crick, Baars, Dehaene, Tononi, Koch e LeDoux, e o paradigma afetivista, onde revisamos as ideias centrais de Panksepp. No segundo capítulo, nós revisamos as ideias centrais de Damásio sobre emoção, sentimento, *self* e consciência, analisando tanto a atualização conceitual e neurobiológica no decorrer de suas obras principais quanto como cada teoria está embasada nas ideias de James sobre os processos emocionais e nos paradigmas discutidos no primeiro capítulo. No terceiro capítulo, nós revisamos as ideias centrais do modelo construtivista da emoção de Russell e Barrett, as posições de Damásio e outros autores sobre a problemática da possibilidade de consciência em animais não-humanos, avaliamos as teorias centrais de Damásio frente ao contexto neurobiológico atual, e analisamos tanto os problemas conceituais inerentes à teoria da consciência de Damásio quanto as posições de Damásio sobre o problema da causação mental e identidade mente-cérebro. Nós esperamos, com essa dissertação, ter contribuído, mesmo que modestamente, com um debate que vem ganhando grande impacto na literatura especializada, que vai muito além do aqui foi apresentado.

Palavras-chave: Consciência; Emoção; *Self*; Sentimento; Damásio; Neurobiologia; Filosofia da Mente.

Abstract

In this dissertation we sought to make a philosophical analysis of Antonio Damasio's theories on emotion, feeling, *self*, and consciousness in the face of the contemporary neurobiological context. In the first chapter, we reviewed the literature on the two paradigms of consciousness in neurobiology, in this case, the cognitive paradigm, where we reviewed the works of Crick, Baars, Dehaene, Tononi, Koch and LeDoux, and the affectivist paradigm, where we reviewed Panksepp's core ideas. In the second chapter we review Damasio's central ideas on emotion, feeling, *self*, and consciousness, analyzing both the conceptual and neurobiological actualization in the course of his major works, and how each theory is based on James' ideas on emotional processes and the paradigms discussed in the first chapter. In the third chapter, we reviewed the core ideas of Russell's and Barrett's constructivist model of emotions, the stance of Damasio and other authors on the problematic about the possibility of consciousness in nonhuman animals, we assessed Damasio's central theories in the face of the current neurobiological context, and we analyzed both the conceptual problems inherent in Damasio's theory of consciousness and Damasio's stance on the problem of mental causation and mind-brain identity. We hope, with this dissertation, to have contributed, even modestly, to a debate that has been gaining great impact in the specialized literature, which goes far beyond what was presented here.

Keywords: Consciousness; Emotion; *Self*; Feeling; Damasio; Neurobiology; Philosophy of Mind.

Sumário

Introdução.....	1
A neurofilosofia de António Damásio.....	5
 Capítulo 1 - Os dois paradigmas centrais da consciência (emocional) nas neurociências.....	8
O paradigma cognitivista de consciência.....	9
Fundamentos cognitivistas.....	9
Crick.....	11
Baars.....	16
Dehaene.....	20
Tononi e Koch.....	29
LeDoux.....	41
A consciência cognitiva.....	47
O paradigma afetivista da consciência.....	48
Panksepp.....	50
A consciência afetiva.....	56
 Capítulo 2 – A teoria da consciência de Damásio: conceitos-chave e fundamentos teóricos.....	58
Corpo e emoção.....	58
A teoria das emoções e sentimentos.....	59
James.....	60
Emoção e sentimento segundo Damásio.....	62

Fundamentos teóricos da teoria da emoção e sentimento de Damásio....	73
A teoria do <i>self</i>	76
Fundamentos da teoria de Damásio sobre <i>self</i>	81
A teoria da consciência.....	82
Fundamentos da teoria de Damásio sobre a consciência.....	86
 Capítulo 3 - Análise dos fundamentos científico e filosóficos da teoria de António Damásio.....	 89
O modelo construtivista de emoção.....	89
A possibilidade de consciência animal não-humana.....	96
A teoria da emoção, <i>self</i> e consciência de Damásio no contexto neurobiológico atual.....	102
A teoria das emoções e sentimentos.....	102
A teoria do <i>self</i>	111
A teoria da consciência.....	112
Os problemas conceituais da teoria das emoções e consciência de António Damásio.....	116
Reduccionismo e poder causal da consciência em Damásio.....	124
 Conclusão.....	 127
 Referências.....	 130

Introdução

O que é consciência? Seria consciência uma propriedade do cérebro? Se sim, quais processos e/ou áreas cerebrais seriam responsáveis pela consciência?

De certo modo, podemos dizer que as questões acima representam o que é comumente chamado de “problema da consciência”, que foi muito bem formalizado por Chalmers (1995, 1996), mesmo que este autor tenha uma postura muito distinta da que será utilizada no decorrer desta dissertação.

Na presente dissertação, toda a argumentação será embasada na proposta neurofilosófica apresentada por Patricia Churchland em seu livro de 1986, *Neurophilosophy*. O propósito dessa autora pode ser colocado da seguinte maneira: apresentar filosofia da ciência para neurocientistas e neurociência para filósofos. Não de uma maneira superficial como vinha sendo feito naquela época, onde era comum usar argumentos reducionistas generalizados e/ou superficiais, como os argumentos de Smart (1959) e Place (1956)¹. Patricia Churchland utilizou, ao invés, teorias contemporâneas daquela época, apresentando tanto o método do pensamento neurocientífico para filósofos quanto as implicações dos métodos e dados neurocientíficos para a filosofia da mente, sempre tendo em vista a importância das discussões e temas da filosofia da ciência para integrar um modelo que é, ao mesmo tempo, neurocientífico (ciência teórica) e filosófico. De modo geral, neurofilosofia pode ser definida do seguinte modo:

Neurofilosofia surge do reconhecimento que, finalmente, as ciências do cérebro e suas tecnologias adjuntas estão suficientemente avançadas a um ponto que um progresso real pode ser feito no entendimento da

¹ Por exemplo, o famoso argumento de que dor é igual ao disparo das fibras C, que, na verdade, são responsáveis por apenas um aspecto da transmissão da dor, chamada de segunda dor, que leva mais tempo para ocorrer (Price, 1972; Dubin & Patapoutian, 2010). Naquela época, a explicação para a pouca acurácia dos argumentos de identidade era que o que realmente importava era a natureza do argumento em si: que uma neurociência madura traria padrões de identidade cada vez mais precisos (Blickle et al., 2012). Já Patricia Churchland (1986), no *Neurophilosophy*, propôs um modelo que também leva em conta a acurácia dos argumentos de identidade, englobando o método de pensamento neurocientífico na argumentação filosófica (filosofia da ciência).

relação mente-cérebro. Mais ousadamente, ela prediz que a filosofia da mente conduzida sem o entendimento sobre os neurônios e o cérebro é provavelmente estéril. Neurofilosofia, conseqüentemente, foca os problemas na intersecção entre uma neurociência jovem e uma filosofia já grisalha. (Churchland, 2002, pág. 3, tradução nossa).

Seguindo esse modelo de raciocínio, o presente trabalho parte do princípio de que a consciência, assim como qualquer outro fenômeno psicológico, possui uma natureza neural, portanto, não teria como abordar esse assunto sem levar em conta os argumentos e modelos neurobiológicos sobre esse tipo de estado mental. Mas antes de partirmos para a discussão central desta dissertação, cremos que seja interessante abordarmos, mesmo que rapidamente, alguns métodos centrais na neurociência cognitiva, afetiva, e comportamental, já que esses procedimentos irão ser mencionados no decorrer dessa dissertação, no caso, as medições de correlação positiva e negativa através de fMRI e PET.

Alguns métodos neurocientíficos

As técnicas de neuroimagem, além de terem sido um enorme avanço para as ciências da mente e do cérebro, também forneceram um novo arcabouço teórico para a filosofia da mente (Armstrong, 1999). A partir dessas técnicas, os argumentos de identidade “atividade cerebral – fenômeno psicológico” se tornaram muito mais fortes e cientificamente embasados. Nas neurociências (cognitiva, afetiva, comportamental e social) as funções de uma determinada área cerebral (tanto cortical quanto subcortical) são dadas a partir de “correlação positiva”, isto é, a função de uma determinada área cerebral é dada através da correlação entre a ativação dessa área e a presença de um determinado processo psicológico.

E, seguindo a mesma lógica, correlações negativas são estabelecidas tanto a partir da ativação de uma determinada área cerebral e a supressão de alguma atividade cognitiva e/ou comportamental (por exemplo, na literatura está estabelecido que a

ativação dos córtices pré-frontais ventromedial e ventrolateral são responsáveis pelo *controle* da ansiedade) quanto a partir de um processo cognitivo e/ou comportamental decorrente da menor atividade de uma determinada área cerebral (seguindo o mesmo exemplo, foi observado que processos intensos de ansiedade estão correlacionados com menor atividade dos córtices pré-frontais ventromedial e ventrolateral) (Etkin e Wager, 2007). E foram as técnicas de neuroimagem as responsáveis pelo estabelecimento dessas correlações.

Imagem por ressonância magnética funcional (fMRI)

A técnica de neuroimagem mais usada por inúmeras vertentes empíricas das neurociências é certamente a fMRI. Em termos gerais, fMRI é uma técnica de neuroimagem responsável por captar as propriedades paramagnéticas das hemoglobinas oxigenadas e desoxigenadas da corrente sanguínea. É por esse motivo que a fMRI é geralmente chamada de “fMRI BOLD”, onde BOLD significa *blood-oxygen-level dependent* (dependente do nível de oxigênio no sangue). Essa técnica é usada para captar áreas cerebrais que estão mais ativadas e/ou desativadas durante determinada atividade cognitiva e/ou comportamento numa resolução de 2 a 3 milímetros, onde, geralmente, maior atividade é representada em escalas de vermelho (incluindo laranja e amarelo), e menor atividade é representada em escalas de azul (incluindo o roxo, lilás e o verde) (Friston et al., 1998). Como pode ser visto na imagem abaixo:

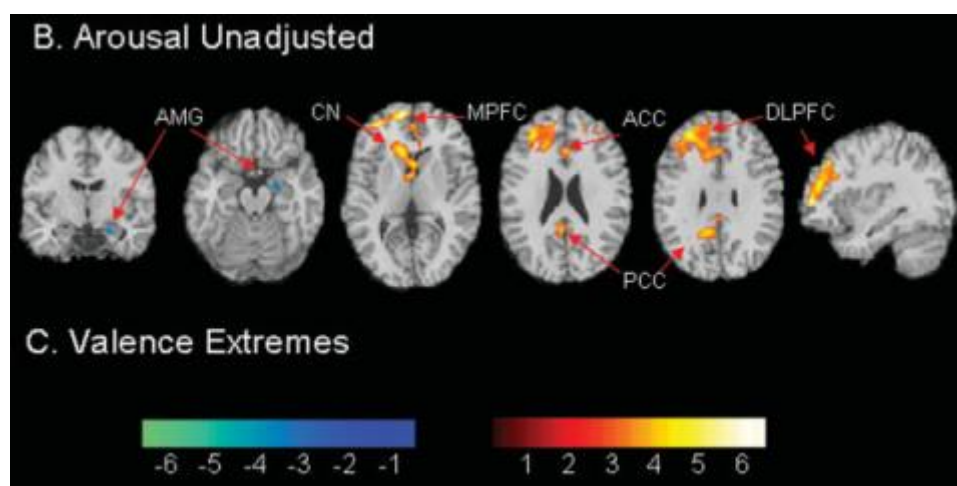


Fig. 1. Padrões de ativação neural capitados por fMRI. Imagem tirada de Posner et al. (2009, pág. 889).

A partir da imagem acima, é possível observar que o processo emocional testado está positivamente correlacionado com algumas áreas pré-frontais, como o córtex anterior cingulado (ACC), o córtex pré-frontal dorsolateral (DLPFC) e o córtex pré-frontal medial (MPFC), ao mesmo tempo em que esse processo está negativamente correlacionado com a atividade da amígdala (AMG).

Tomografia por emissão de pósitrons (PET)

PET é responsável por captar a radiação de elementos químicos injetados no sangue, chamados de *radiotraçadores*, quando estes chegam ao cérebro através da corrente sanguínea, produzindo tanto imagens 2D quanto 3D a partir da distribuição desses elementos químicos no cérebro. A vantagem de se usar PET é que essa técnica permite observar tanto o fluxo de sangue e oxigênio quanto o metabolismo de glutamato realizado por células nervosas, de modo que o PET se tornou o melhor método para captar imagens do funcionamento cerebral, tanto em termos de resolução quanto de velocidade (dentro de 30 segundos). O padrões de cor para interpretar os dados de correlação usados a partir dessa técnica é semelhante ao usado a partir do fMRI, onde os tons em escala de vermelho representam picos de ativação enquanto tons da escala de

roxo representam padrões de desativação (Grafton et al., 1992), como veremos no segundo capítulo, quando discutirmos a teoria das emoções de Damásio.

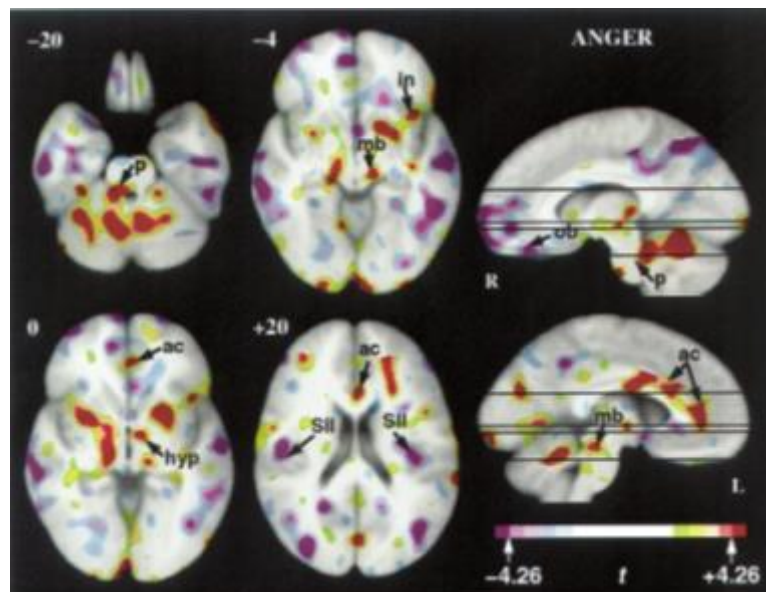


Fig. 2. Padrões de ativação neural capitados por PET. Imagem tirada de Damásio et al. (2000, pág. 1051).

A neurofilosofia de António Damásio

Como dito anteriormente, a partir do ponto de vista que todo processo psicológico possui uma natureza neurobiológica, não seria possível discutir qualquer um desses fenômenos sem levar em consideração dados neurocientíficos sobre o mesmo. Para tanto, nesta dissertação, escolhemos como fio-condutor para a discussão sobre a neurofilosofia da consciência e processos afetivos, a rica obra de António Damásio (1994, 1999, 2003, 2010).

Essa escolha se baseia no fato de que a obra de Damásio, por natureza, dialoga constantemente com a filosofia. Os livros e os artigos teóricos de Damásio não consistem de uma descrição pura de achados empíricos. Na verdade, o neurocientista interpreta os dados apresentados de modo filosófico, indo muito além da mera descrição dos experimentos e achados empíricos sobre determinado fenômeno. Muitas vezes, Damásio até adota uma narrativa que se aproxima da literatura, descrevendo os

processos mentais de uma forma um tanto quanto poética. De um jeito ou de outro, Damásio nunca perde de vista a importância da fundamentação empírica de suas ideias, mesmo quando o autor propõe interpretações que vão além do conteúdo restrito dos dados empíricos que as baseiam.

Outras características importantes da obra de Damásio devem ser acentuadas. Uma delas, como veremos posteriormente, é que o neurocientista coloca o corpo como um objeto central de sua teoria. Mesmo que Damásio não chegue a utilizar expressões como “cognição incorporada” diretamente em sua obra, Damásio deixa claro que não haveria consciência, emoção e sentimento sem corpo. O autor, por exemplo, chega a criticar diretamente a ideia funcionalista clássica que seria possível haver cognição, consciência e sentimento sem corpo (Churchland, 1984/2013). É também interessante notar que Damásio define consciência como sentimento, abraçando uma posição fenomênica ao invés de puramente empírica.

Em termos gerais, a presente dissertação será dividida em três capítulos centrais. No primeiro capítulo, nós analisaremos os dois grandes paradigmas da consciência nas neurociências, nos quais a teoria de Damásio se baseia, no caso, o paradigma afetivista e o paradigma cognitivista. Como poderemos ver no segundo capítulo, a teoria de Damásio não dialoga com apenas um dos dois modelos neurobiológicos da consciência (emocional) – cognitivista ou afetivista –, pois mesmo que Damásio de certo modo baseie sua teoria das emoções na visão de James sobre a mediação do corpo e defenda que a consciência é antes de tudo algo sentido, o que é similar à postura de autores da neurociência afetiva, como Panksepp e Zajonc, Damásio não perde de vista a importância de processos cognitivos para a construção da consciência, inclusive fazendo uso de termos da neurociência cognitiva e ciência cognitiva, como representação e imagem (além de defender um modelo cortical de consciência ampliada, como veremos

posteriormente). De todo modo, ficará claro que a análise das teorias da consciência dentro do paradigma cognitivista é muito mais extensa que a do paradigma afetivista, já que a visão cognitivista de consciência é a dominante no contexto neurobiológico atual (LeDoux. 2015).

No segundo capítulo, nós analisaremos a teoria das emoções, *self* e consciência de Damásio, apresentando as mudanças conceituais e empíricas de seus conceitos centrais, no caso, emoção, sentimento, *self* e consciência (já que, como veremos, há uma conexão direta entre emoção, sentimento, *self* e consciência na obra de Damásio), além de analisarmos a influência do pensamento de James (1884) no modelo de emoção e sentimento do autor. Também analisaremos como que os paradigmas apresentados no primeiro capítulo influenciam na construção dos conceitos e ideias centrais de Damásio.

E, no terceiro capítulo, nós analisaremos as ideias de Damásio frente o paradigma cognitivista dominante atual, além de alguns problemas conceituais inerentes ao seu modelo de consciência, *self*, emoção e sentimento. Como veremos nesse capítulo, as ideias de Damásio sofrem alguns problemas empíricos frente a alguns modelos cognitivistas vigentes da emoção, baseados na ideia de que esse fenômeno é provavelmente algo construído, não sendo nem inato nem vinculado aos circuitos de sobrevivência. Além disso, analisaremos como que o modelo conceitual de Damásio dá origem a inúmeras dificuldades teóricas, devido a uma falta de rigor conceitual de sua teoria, o que também causa problemas teóricos em sua posição sobre reducionismo e causação mental.

Capítulo 1

Os dois paradigmas centrais da consciência (emocional) nas neurociências

Como todo fenômeno psicológico, a discussão sobre as bases neurobiológicas da consciência (perceptiva) e da consciência emocional depende do que é considerado como consciência e emoção antes de tudo, de modo que a interpretação dos dados depende diretamente dos conceitos adotados pelos pesquisadores. Nesse capítulo, nós discutiremos as duas grandes vertentes da consciência (e consciência emocional) na neurobiologia contemporânea: o paradigma cognitivista de consciência e o paradigma afetivista de consciência. Cada um desses grupos possui modelos de consciência, emoção e sentimento distintos baseados diretamente em como cada paradigma conceitua e identifica cada um desses processos, de modo que os dados empíricos usados para discutir e formalizar o modelo de consciência e/ou emoção depende do modelo teórico base de cada paradigma conceitual.

Como será possível observar, o paradigma cognitivista possui uma gama muito maior de autores de grande representatividade em relação ao paradigma afetivista, por ser, certamente, o paradigma mais aceito e publicado nas grandes revistas da área². De qualquer forma, é importante ressaltar aqui que nós apenas analisaremos cada teoria cognitivista brevemente, não é nosso objetivo fazer uma análise aprofundada de cada uma delas, já que isso fugiria do escopo da presente dissertação. O objetivo central aqui é apresentar as ideias gerais das teorias da consciência mais influentes e/ou inovativas atualmente, de modo a apresentar o contexto neurobiológico geral sobre o assunto.

² Por exemplo, a importante revista *Nature Reviews Neuroscience* (a revista com o maior fator de impacto em neurociências) publicou 53 revisões (*reviews*) relacionadas com consciência, todas dentro do paradigma cognitivista da consciência.

O paradigma cognitivista de consciência

Aqui apresentaremos alguns dos modelos teóricos cognitivistas da consciência que alcançaram ou têm alcançado grande impacto na literatura, no caso, as teorias de Crick (e Koch), Baars, Dehaene, Tononi e Koch, e LeDoux. Mas, antes de analisarmos alguns modelos de consciência dentro do paradigma cognitivista, é necessário apresentar os fundamentos desse modelo teórico de mente.

Fundamentos cognitivistas

Segundo Thagard (2014), a hipótese central da ciência cognitiva é que mente pode ser entendida em termos representacionais e computacionais. E isso é visível em qualquer artigo sobre a concepção cognitivista de mente e/ou consciência, tanto que se tornou comum utilizar termos computacionais para se referir à atividade cerebral, como processamento, computação (de informação), codificação, decodificação, análise, e, o mais comum, representação. A tese principal desse paradigma é que o cérebro decodifica e representa informação ambiental. Na verdade, para ser mais preciso, a tese principal da neurociência cognitiva é que *neurônios* decodificam e representam informações ambientais (o modelo neuronal de disparo/não-disparo é próximo do modelo binário da programação computacional (1 e 0)) (Fresco, 2012).

De modo geral, o modelo cognitivista de mente está aliado com a visão de que a mente é formada por um vasto conjunto de representações mentais (de forma análoga aos algoritmos computacionais). Como posto por Thagard (2014), “os teóricos cognitivistas têm proposto que a mente contém tais representações mentais como proposições, regras, conceitos, imagens e analogias lógicas, e que ela usa procedimentos

mentais, como dedução, busca, correspondência, rotação e recuperação” (Thagard, 2014, página de web, tradução nossa).

Em termos gerais, podemos dizer que a cognição é inerente ao modelo representacionalista de mente, de modo que a teorização sobre qualquer processo cognitivo, como conhecimento, atenção, memória, julgamento, avaliação, resolução de problemas, tomada de decisão, compreensão e linguagem, está baseada em modelos representacionais. E, como se sabe, todos esses processos cognitivos se dão a partir da interação com o ambiente, ou seja, são aprendidos [ou condicionados, como argumenta Panksepp e Solms em Panksepp et al., 2016], ou seja, os processos cognitivos só se originam através da interação entre organismo e ambiente (a linguagem, por exemplo, deve ser aprendida através da interação entre organismo e ambiente social). E, como veremos na próxima seção, Panksepp, por exemplo, considera que processos afetivos, ao contrário dos cognitivos, são inatos (incondicionados). É importante manter essa distinção em mente.

De modo geral, o paradigma cognitivista da consciência é caracterizado por duas ideias centrais: que a consciência é um fenômeno cognitivo (dado por um conjunto de representações mentais) e que a consciência é um fenômeno cortical. E uma ideia dialoga diretamente com a outra.

O trabalho de Penfield e Boldrey (1937), onde os autores usaram estimulação elétrica para investigar representações corticais de natureza motora e sensorial, continua sendo um artigo com grande impacto na área³. O que é importante desse artigo para nossa análise aqui é que esses autores demonstraram que a representação de informação sensorial é realizada quase que exclusivamente por áreas corticais. Mesmo que se possa argumentar que áreas subcorticais, como o tálamo, sejam fundamentais para isso, já que

³ Até o momento, foi citado 3082 vezes.

o tálamo é responsável por realizar uma pré-avaliação do estímulo sensorial antes de enviá-lo para a área cortical correspondente, como a pré-avaliação de informação visual pelo tálamo visual antes de enviá-la ao córtex visual (LeDoux, 1994, 2015), é importante ressaltar que essa pré-avaliação é muito mais básica (menos detalhada) do que o processamento cortical do estímulo. Seguindo o exemplo “galho ou cobra” (ver pág. 110 desta dissertação), o tálamo visual interpreta um galho como um estímulo potencialmente nocivo com base em informação inata sobre o formato de uma cobra, apenas após o processamento do estímulo pelo córtex visual que é possível reconhecer que o objeto é, na verdade, um galho. Portanto, representações corticais detalhadas são atribuídas às áreas corticais.

Logo, se consideramos que são as áreas corticais as responsáveis por representarem informações sensoriais e que a consciência se baseia em tais representações sensoriais (como experiência visual, auditiva, tátil, etc.), representações corticais são o suficiente para dar origem à consciência, ainda mais se levarmos em consideração que as áreas pré-frontais possuem conexões com todas as áreas sensoriais (Dehaene e Changeux, 2011, pág. 207). Mesmo propriocepção e interocepção são processadas pelo córtex somatossensorial (London et al., 2008) e pela ínsula (Craig, 2002, 2003, 2009; Damasio, 2010; Barrett e Simons, 2015) respectivamente.

Portanto, além de representacionalista, o paradigma cognitivista é fundamentalmente corticalista. E a importância das áreas corticais, principalmente das áreas pré-frontais, data o início da investigação neurobiológica da consciência proposta por Crick e Koch.

Crick

As hipóteses de Crick (e Koch)⁴ podem ser consideradas relativamente, mas não totalmente, desatualizadas hoje em dia, porém cremos que poucos negariam a grande importância de Crick para o desenvolvimento da pesquisa neurocientífica da consciência, tanto que uma das primeiras grandes afirmações de seu famoso artigo de 1990 é sobre um certo menosprezo da comunidade neurocientífica em relação ao problema da consciência naquela época (Crick e Koch, 1990). De qualquer forma, não podemos menosprezar a grande importância de Crick e Koch para a propulsão dessa temática na neurobiologia, e, por esse motivo, cremos que seja necessária uma breve análise desse artigo de 1990.

Uma das primeiras afirmações desse artigo foi que uma definição de consciência, naquele estágio do desenvolvimento científico, seria algo prematuro a se fazer⁵. E é importante ressaltar que Crick mantém essa posição mesmo nos artigos subsequentes (Crick e Koch, 1998, 2003), uma vez que o interesse principal do autor era avaliar e propor o estudo da consciência a partir de correlatos empíricos, como EEG e MRI, ou seja, a partir de modelos estritamente experimentais. Para o autor, uma definição de consciência só poderia ser feita após um grande acúmulo de evidências sobre os correlatos neurais desse fenômeno, de modo que fosse possível distingui-lo de outros processos biológicos. Desse modo, todas as hipóteses do autor são puramente empíricas, utilizando-se de modelos teórico-experimentais para o entendimento da consciência.

Outra posição importante desse artigo é o argumento que a discussão sobre consciência em outros animais, principalmente em não-primatas, não é frutífera.

⁴ Crick e Koch, na verdade, mas aqui preferimos optar pela distinção entre os autores já que, hoje em dia, as posições de Koch sobre o problema da consciência são consideravelmente distintas daquela época.

⁵ Como posto pelos próprios autores: “todo mundo tem uma ideia aproximada do que se entende por consciência. Sentimos que é melhor evitar uma definição precisa de consciência por causa dos perigos de uma definição prematura. Até que entendamos melhor o problema, qualquer tentativa de uma definição formal provavelmente será enganosa ou excessivamente restritiva, ou ambas” (Crick e Koch, 1990, pág. 264, tradução nossa).

LeDoux (2015), por exemplo, discute brevemente sobre essa questão. Segundo este autor, propor que animais não-humanos possuem graus de consciência ainda continua fora do estágio atual da neurobiologia, e talvez nunca venha a ser possível apresentar evidências que comprovem tal hipótese, pois, querendo ou não, a consciência ainda é estudada com base em relatos pessoais, o que necessariamente demanda *linguagem* (LeDoux, 2015).

Outro ponto também importante é o argumento de que, segundo Crick e Koch (1990, pág. 264), ainda não temos condições de compreender a natureza ou verificar se realmente existe o que os filósofos chamam de *qualia*, algo também discutido recentemente por LeDoux (2015) e Dehaene (2014), já que, segundo alguns filósofos, como Jackson (1982, pág. 273), os *qualia* estão fora no alcance da experimentação física.

De qualquer modo, a contribuição mais famosa do artigo de Crick e Koch (1990, pág. 271 e 272) é a hipótese sobre as oscilações de 40 Hz (ondas gama) para explicar a sensação de *binding* da experiência consciente. A hipótese sobre oscilações neuronais para explicar o *binding* da experiência perceptual já estava presente na literatura antes da publicação do artigo de 1990, como o modelo de reconhecimento visual de Milner (1974)⁶. E pouco tempo depois da publicação de Crick e Koch (1990), Llinás e Ribary (1993) propuseram evidências que demonstraram que oscilações de 40 Hz também estão presentes no sono REM, comumente correlacionado com sonhos (atividades conscientes). Mas foi através da publicação de Crick e Koch (1990) que essa hipótese ganhou força, se tornando sinônimo de atividade consciente em trabalhos empíricos posteriores.

⁶ Basicamente, Milner (1974) hipotetizou que o reconhecimento visual, como perceber se um objeto está se movendo para a esquerda ou para a direita do observador, seria dado através de sincronização de disparo neuronal durante um determinado intervalo de tempo (sincronização neuronal), o que pode ser observado através de ondas gamas por eletroencefalografia, como feito por Llinás e Ribary (1993).

A hipótese das oscilações de 40 Hz mesmo já não sendo tão forte, não foi completamente descartada. Dehaene (2014) argumenta que mesmo que essas oscilações, assim como outros níveis de oscilação gama, também estejam presentes em atividades neurais não-conscientes (Aru et al., 2012), esse tipo de oscilação se torna mais forte quando um determinado estímulo se torna consciente. Ou seja, segundo Dehaene (2014), não é apenas a presença de ondas gama que indica atividade consciente, mas sim seu incremento durante atividade neural consciente (Whittingstall e Logothetis, 2009), geralmente 300 ms após a apresentação do estímulo (Gaillard et al., 2009; Dehaene, 2014). Desse modo, a hipótese de Crick não estaria incorreta, apenas incompleta devido ao desenvolvimento empírico da época (de qualquer forma, ainda está em aberto se esse tipo de onda realmente pode ser tido como um correlato neural da consciência realmente confiável, como discuto por Koch et al., 2016, pág. 313).

A hipótese da oscilação de 40 Hz de Crick e Koch (1990) realmente foi muito importante para a época e continua sendo, mas nós também gostaríamos de chamar atenção para uma contribuição que tomou conta e continua sendo a posição mais aceita dentro do cenário científico sobre a consciência: a importância das áreas corticais, principalmente das áreas frontais, para a criação da consciência. No artigo de 1995, Crick e Koch argumentam que o córtex visual primário (V1) não seria o responsável pela consciência visual por ser de baixa hierarquia no processamento visual e por possuir pouquíssimas conexões com outras áreas cerebrais, como o córtex pré-frontal (Crick e Koch, 1995, pág. 122). Algo que é reafirmado pelos autores no artigo 1998, onde eles argumentam que o modelo de consciência visual seria o melhor modo para se descobrir as bases neurais da consciência (Crick e Koch, 1998, pág. 100 e 103). A nosso ver, essa foi uma das primeiras hipóteses a sugerir tanto uma interconexão entre diversas áreas cerebrais quanto uma certa posição privilegiada do córtex pré-frontal na

criação da consciência, que é hoje a posição predominante dentre as hipóteses neurobiológicas sobre esse fenômeno⁷ (Dehaene, 2014; LeDoux, 2015; Koch et al., 2016), tanto que poucos autores defendem bases subcorticais da consciência [Damásio é um deles, como veremos no segundo capítulo].

Em seu último artigo de grande importância, cremos que a maior contribuição, e que permanece até hoje, é a tese da necessidade de o estímulo visual ser processado por camadas corticais de alta ordem e, em seguida, o resultado desse processamento deve ser enviado ao córtex pré-frontal para se tornar consciente (Crick e Koch, 2003, pág. 121). Podemos dizer que a versão contemporânea dessa hipótese é o modelo de “broadcasting” proposto por Baars (1988), que é o modelo mais defendido atualmente (Dehaene e Naccache, 2001; Dehaene, 2014; LeDoux, 2015), o qual abordaremos adiante.

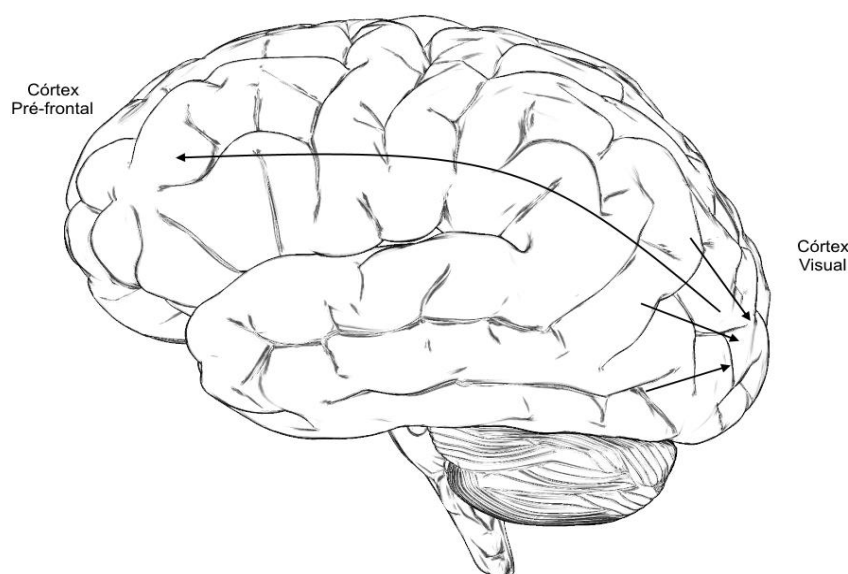


Fig. 3. Ilustração nossa baseada no modelo proposto por Crick e Koch (2003), onde o estímulo visual é processado por níveis superiores do córtex visual para depois ser transmitido para o córtex pré-frontal (imagem de nossa autoria).

⁷ Para se ter uma ideia geral, o artigo mais citado de Panksepp foi publicado em 2005 e recebeu, até o momento (12 anos), 831 citações segundo o Google Scholar. Já o artigo de Dehaene e Changeux de 2011 (ou seja, mesmo sendo mais recente) foi citado 946 vezes até a escrita dessa dissertação (intervalo de 6 anos). E esse nem foi o artigo mais citado de Dehaene, que é, na verdade, o artigo de Dehaene e Naccache de 2001, citado 1823 vezes até o momento.

O importante aqui é perceber tanto a importância histórica dos artigos de Crick e Koch (1990, 1995, 1998, 2005), através dos quais a consciência se tornou um assunto de relevância na neurobiologia, quanto a importância de algumas ideias teórico-experimentais que continuam relevantes para os dias atuais, no caso, a observação de atividade consciente a partir de ondas gama e a importância das áreas corticais (principalmente as pré-frontais) na sustentação da consciência.

Mas, diferentemente do modelo empírico e neurobiológico de Crick, os modelos cognitivos nem sempre estabelecem uma conexão precisa entre consciência e neurobiologia, utilizando-se, ao invés, de modelos computacionais para refletir sobre a atividade consciente, como o modelo do proeminente autor, Bernard Baars.

Baars

Cronologicamente falando, a hipótese de Baars (1988) sobre o espaço de trabalho global foi publicada antes das hipóteses de Crick, porém esse modelo foi majoritariamente embasado no cognitivismo e computacionalismo da época, possuindo pouco embasamento neurobiológico (a fundamentação neurobiológica dessa teoria só ocorreu posteriormente com as publicações dos artigos de Dehaene, como veremos adiante). Analisar esse modelo aqui, mesmo que brevemente, é de grande importância, uma vez que ele é, atualmente, o modelo mais aceito para a explicação da consciência (LeDoux, 2015). E é importante ressaltar que Baars, assim como Crick, publicou apenas trabalhos teóricos sobre a consciência, porém, diferentemente de Crick, que focou mais sobre as bases neurobiológicas da consciência, Baars focou e continua focando majoritariamente sobre as bases cognitivistas e computacionais desse fenômeno, algo

que Crick e Koch consideraram um método pouco factível para descobrir as reais bases da consciência (Crick e Koch, 1990, pág. 264 e 265).

A teoria de espaço de trabalho global (GWT) foi baseada em modelos computacionais propostos por Alan Newell e Herbert Simon (Baars, 1988, 2002), porém esses autores não fizeram nenhuma correlação com o funcionamento da consciência, o que só foi proposta por Baars em 1988. É pelo espaço de trabalho global (GW) que um processamento cerebral inconsciente torna-se consciente. Segundo a teoria, há inúmeros processamentos informacionais inconscientes sendo realizados no cérebro nesse exato momento, que competem entre si para atingir o GW⁸, onde a atividade é então aumentada⁹, se espalhando para outras áreas cerebrais para então retornar ao GW e assim sucessivamente, ou seja, esse modelo é, além de tudo, um modelo de integração de informação (Baars, 1998, 2005).

Esse sistema é também a base do modelo atencional de Baars (1998, 2005). Para explicar o processo atencional, Baars utiliza o que ele chama de “metáfora do teatro”. Nessa metáfora o autor propõe que a atenção é o foco de luz em um teatro enquanto o resto do ambiente permanece nas sombras. As múltiplas atividades sensoriais competem para receber o foco da atenção e se tornar conscientes. Literalmente, de acordo com a metáfora, elas competem para ficar sob o *spotlight* da atenção por um determinado período de tempo. De acordo com o autor, esse ciclo ocorre aproximadamente a cada 100 ms (Baars e Franklin, 2003).

⁸ Hipótese similar à apresentada por Dennett (1991), onde narrativas fragmentadas (informação) competem entre si para se tornar conscientes a partir de um mecanismo chamado de *pandemonium*, tanto que, em 2001, o próprio autor toma a GWT como a melhor explicação para a consciência (Dennett, 2001).

⁹ Informação importante para quando formos apresentar a teoria de Dehaene.

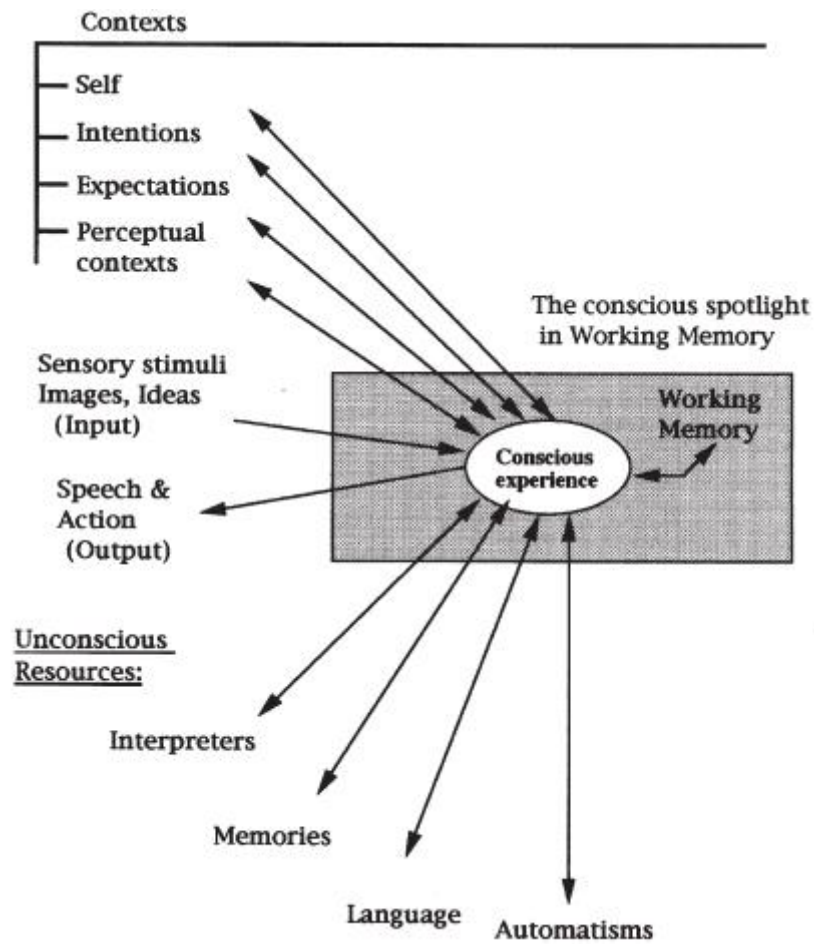


Fig. 4. Modelo para o GW e metáfora do teatro de Baars (1998).

A partir da ilustração acima, podemos perceber a inter-relação entre memória de trabalho e o GW (consciência). Memória de trabalho é o uso da memória de curto e longo prazo [esta última através do *episodic buffer* (Baddeley, 2000)] em tarefas cognitivas (*cognitive tasks*). Seguindo esse modelo, o processamento sensorial consciente possui a capacidade de ser transmitido a inúmeras áreas cerebrais, algo que não ocorre com o processamento inconsciente, e, além disso, através do GW inúmeros processamentos neurais distintos são momentaneamente ligados (*binding*), dando origem, por exemplo, a uma sensação de *self*, que inclui inúmeros processos distintos, como memória, linguagem, intenções, sensações corporais, etc.. Baars (2002) sumariza o modelo acima do seguinte modo: a percepção consciente é mais do que apenas

processamento perceptual, uma vez que este é restrito às próprias regiões cerebrais responsáveis por essa função, sendo majoritariamente inconsciente; é pelo GW (consciência) que é possível compreender informações novas, principalmente as que exigem processamento linguístico; a memória de trabalho depende de inúmeros processos cognitivos distintos, como percepção, discurso interior (*inner speech*) e imagens visuais (criadas ou presentes no ambiente); o controle voluntário baseado em metas conscientes e percepção de resultados [redundância]¹⁰; a atenção seletiva propicia acesso ao conteúdo consciente e vice-versa [outra redundância]¹¹; e a consciência possibilita o acesso ao *self* [que é em si consciente na teoria do autor, ou seja, outra redundância]¹².

Velmans (2009, pág. 73) faz um ótimo resumo sobre as principais funções da consciência segundo a GWT.

- É pela consciência que um *input* é definido como tal, ou seja, só é *input* quando a informação de entrada se torna consciente (isto é, só tomamos ciência de um estímulo quando este se torna consciente).
- A consciência é necessária para resolução de problemas e aprendizado, principalmente quando se envolve novidade. Isso é discutido pelo próprio Baars (1997), argumentando que não haveria aprendizado sem consciência, já que sem a integração de informação disponibilizada pelo GW só existiria processamentos inconscientes e distintos no cérebro.

¹⁰ O que gera certa redundância, já que, segundo a própria teoria, o controle voluntário é sempre consciente, logo, é o mesmo que dizer que um processo consciente precisa de consciência. Desse modo, Baars propõe um modelo pelo qual ele define um conceito utilizando o próprio conceito que quer definir. Por exemplo, autor define atenção consciente através do conceito de consciência (Baars, 1988, pág. 33), ao mesmo tempo em que seu próprio modelo de consciência depende do conceito de atenção seletiva (que é intrinsecamente consciente) para ser formalizado.

¹¹ Atenção seletiva é uma função executiva segundo o cognitivismo, que é sempre consciente.

¹² O *self*, na teoria de Baars, é intrinsecamente consciente (Baars, 1988, pág. 241). Logo, ao se afirmar que a consciência permite acesso a um fenômeno consciente, além de ser redundante, não explica como o *self* é constituído e nem a relação entre *self* e consciência (tomando que são conceitos distintos, o que não é claro na obra de Baars).

- É pela consciência que é possível se adaptar a uma nova situação (integrar novas informações ambientais).
- Já que o GW (consciência) integra a grande quantidade de informação neural, é a partir desse mecanismo que é possível o controle voluntário da ação, utilizando-se de informação motora, perceptual e mnemônica (baseado em experiências e ações anteriores).
- Discurso interior e imaginação consciente possibilitam a reflexão e controle tanto de processos conscientes quanto inconscientes.
- E a consciência (GW) é indispensável em situações que demandam processamento de informações (*inputs*) imprevisíveis, possibilitando respostas flexíveis.

Como podemos observar, o modelo de Baars é muito mais cognitivista e computacionalista do que puramente neurobiológico, mesmo assim é o modelo mais utilizado tanto pela neurobiologia quanto pela robótica nas pesquisas empíricas sobre a consciência (Dehaene, 2014; LeDoux, 2015). Desse modo, a GWT é provavelmente a teoria mais importante sobre esse fenômeno atualmente (Dehaene, 2014; LeDoux, 2015). E sobre a aplicação desse modelo em pesquisas empíricas na neurobiologia, o autor que mais se destaca é Stanislas Dehaene.

Dehaene

Independentemente da posição filosófica, é inegável o fato de que a publicação do artigo de Dehaene e Naccache em 2001, *Towards a cognitive neuroscience of consciousness*, foi um marco na história do entendimento científico da consciência (Dehaene, 2014). Nesse artigo não só Dehaene embasou a GWT com dados neurocientíficos como também trouxe novas ideias para a teoria, criando o então

denominado *Global Neuronal Workspace Model* (GNW)¹³. Como visto acima, a teoria de Baars carecia de evidências advindas da neurobiologia, a GWT era apenas uma hipótese cognitivista e computacional sobre a consciência da qual o próprio autor admitiu não ter tido muito apelo na época (Baars, 2002). O grande objetivo do artigo de Dehaene e Naccache (2001) foi, então, embasar a GWT com dados empíricos.

Dehaene e Naccache (2001) analisaram inúmeras evidências sobre o processamento consciente e inconsciente, tanto em pacientes saudáveis quanto naqueles com alguma patologia que comprometia o funcionamento normal da atividade consciente, resultando nas seguintes observações:

- Inúmeros experimentos, principalmente em pacientes com algum transtorno neurológico, demonstram que há processamento cognitivo sem a necessidade de consciência, como nos casos da agnosia visual (incapacidade de reconhecer ou identificar objetos mesmo sem perdas nas funções sensoriais), prosopagnosia (prejuízos na capacidade de reconhecer faces); acromatopsia¹⁴ (perda da capacidade de reconhecer cores); desconexão do corpo caloso, afasia, amnésia e heminegligência (perdas na capacidade de processar informações sensoriais do hemisfério cerebral oposto). Por exemplo, em pacientes com cérebro dividido (*split-brain patients*)¹⁵, o participante pode relatar não experienciar o estímulo mesmo conseguindo falar o nome do objeto e/ou desenhá-lo, o mesmo ocorre

¹³ É importante que notar o viés *neuronal* do nome.

¹⁴ Alguém poderia questionar sobre a relação entre acromatopsia e *blindsight*, já que no último mesmo o indivíduo não tendo a experiência consciente do objeto, ele consegue desenhá-lo ou falar sobre ele, ou seja, ele processa cognitivamente o objeto, diferentemente do primeiro, onde o indivíduo não possui experiência de cor e não se comporta mediante o processamento cognitivo de cor. Entretanto, o caso de acromatopsia comprova que o processamento cognitivo visual é dividido em inúmeras partes, pois mesmo que o participante não tenha experiência de cor, ele continua tendo a experiência de tamanho, forma, movimento, tonalidade (em escalas de cinza), etc., e que consciência visual se caracteriza pela ligação (*binding*) de inúmeros processos visuais processados por distintas áreas do córtex visual (Celesia, 2010).

¹⁵ Pacientes com cérebros divididos possuem características similares com os pacientes de visão cega. Nos experimentos com esse tipo de paciente, se o objeto é posto no campo visual esquerdo, o paciente não possui *consciência declarativa* sobre o objeto. E se o objeto é posto no campo visual direito, o paciente não consegue falar sobre o objeto por não possuir *consciência visual declarativa* do objeto (Gazzaniga, 1967).

nos experimentos com *blindsight*. [E não podemos nos esquecer dos famosos experimentos sobre percepção não-consciente apresentados por Marcel (1983)¹⁶].

- Evidências demonstram que não há consciência sem atenção. Pacientes com heminigligência direita (lembrando que o hemisfério direito processa informações advindas do campo perceptual esquerdo do organismo) não conseguem perceber estímulos no campo visual esquerdo quando há competição de estímulos, entretanto, quando não há competição, o estímulo no campo esquerdo é percebido pelo participante. Ou seja, o participante não ter visto e nem demonstrado nenhuma atividade consciente em relação ao estímulo (nenhum tipo de relato sobre o estímulo) no campo visual esquerdo quando há competição entre estímulos indica que, na verdade, tal condição neurológica impede que o paciente tenha atividade atencional em relação ao estímulo. Como bem analisado por Koch e Tsuchiya (2006), atenção e consciência, mesmo sendo fenômenos proximamente relacionados, são fenômenos distintos, podendo haver processamento atencional sem consciência, mas nunca consciência sem atenção, mesmo que seja uma atenção quase nula, como a que ocorre para estímulos periféricos. E os autores ainda argumentam que o controle *top-down* da atenção perde muita força quando não há competição entre estímulos, logo, no exemplo acima, o que ocorre é o aumento do controle *top-down* da atenção apenas a um estímulo quando há competição, sendo assim, o outro estímulo se torna completamente inconsciente para o participante. E é interessante notar que

¹⁶ Em termos gerais, Marcel (1983) conduziu 5 experimentos para investigar a relação entre consciência (mais especificamente, consciência de acesso ou consciência declarativa) e processamento de informação linguística visual (palavras escritas) usando a técnica de “mascaramento” de estímulo. De modo geral, os participantes tinham de relatar possíveis relações linguísticas entre o estímulo (a palavra) consciente e o padrão mascarado (palavra mascarada) não-consciente. Já que os participantes conseguiam estabelecer relações linguísticas entre o estímulo conscientemente percebido e o padrão mascarado não percebido conscientemente, Marcel (1983) demonstrou, pela primeira vez na psicologia, que o processamento perceptual (que é um processamento cognitivo) ocorre inconscientemente.

heminigligência direita ocorre devido a lesões no córtex parietal, que é tido como responsável tanto pela consciência (como discutiremos adiante) (Lou et al., 2004) quanto pela atenção (Bushnell et al., 1981; Wojciulik e Kanwisher, 1999; Ester et al., 2016a, 2016b).

- Outra função desse artigo foi embasar empiricamente as funções da consciência, algumas delas já teorizadas por Baars em 1988: a consciência é necessária apenas para algumas operações mentais, como manutenção e explicitação da informação, onde os autores discutem dados que sugerem que a representação visual de um conjunto de letras rapidamente decai a um nível indetectável numa situação em que não há aplicação da atividade consciente; novas combinações de operações, como nos experimentos baseados no efeito Stroop¹⁷; e comportamento intencional, já que o participante deve utilizar experiências passadas para tentar prever as consequências de ações futuras.

Todas as análises empírico-teóricas apresentadas nesse artigo são importantes e deram origem a novas linhas teóricas para o GNW (*global neuronal workspace model*), mas o ponto principal do artigo de Dehaene e Naccache (2001), algo que já tinha sido proposto por Dehaene num artigo anterior (Dehaene et al., 1998), é definir o GW a partir de conexões corticais (cortico-corticais e tálamo-corticais) de longa distância. Esse modelo é hoje considerado praticamente sinônimo de atividade consciente. Como visto anteriormente, Crick e Koch (2003) também propuseram que um estímulo visual

¹⁷ Os experimentos de Stroop (1935) (vastamente reconhecidos na psicologia) são baseados em um procedimento um tanto quanto simples: trata-se de apresentar ao participante o nome de uma determinada cor (estímulo linguístico escrito) em uma coloração distinta da qual o estímulo linguístico se refere, por exemplo, apresentar a palavra “verde” só que com coloração vermelha. Desse modo, pede-se para o participante dizer a palavra escrita ignorando a cor da qual ela é apresentada. Merikle et al. (1995) usando esse mesmo procedimento, observaram que no experimento em que 75% das tentativas eram pareamentos incongruentes conscientes (cor e palavra distintas) os participantes começaram a processar os pareamentos incongruentes mais rapidamente que os pareamentos congruentes (cor e palavra iguais).

só se torna consciente quando este atinge o córtex pré-frontal¹⁸. Mas a sistematização desse modelo só foi proposta por Dehaene em 1998 e aprofundada em 2001. A partir das ideias propostas por Baars (1988), Dehaene e colegas (1998) deduziram que, se o GW realmente existe, é necessário que ele receba conexões de diferentes áreas cerebrais, e para tanto é necessário existir conexões de longa distância. Isso foi representado pelos autores no modelo abaixo:

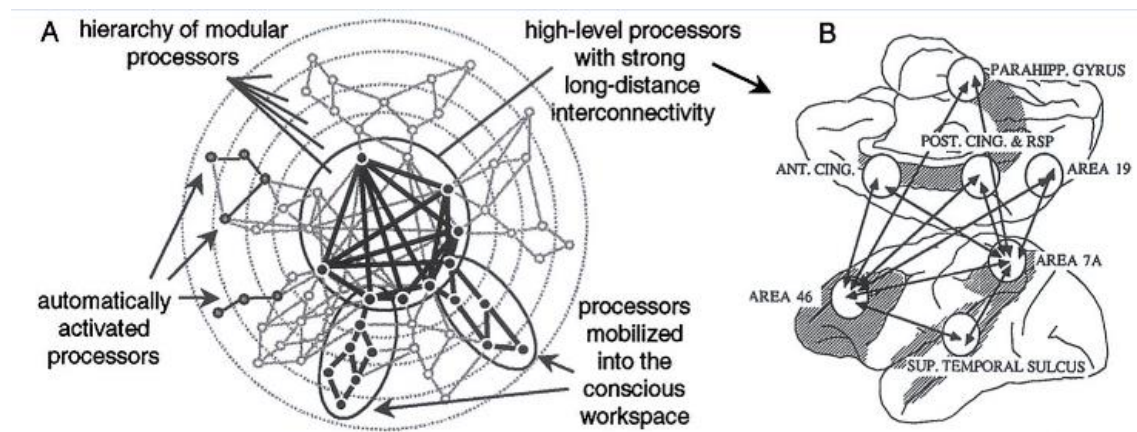


Fig. 5. Nesse modelo podemos perceber que há uma hierarquia entre as conexões de diferentes áreas cerebrais, e bem ao centro se encontra as conexões de alta-ordem, que, segundo os autores, são estabelecidas entre os córtices pré-frontal, parietal, temporal e cingulado anterior, sendo, portanto, o GW neural. (Ilustração tirada de Dehaene e Naccache, 2001, pág. 27).

E é importante notar que, para Dehaene [e a maioria dos autores atualmente], a consciência é instanciada por áreas corticais, principalmente pelos córtices pré-frontais e parietais, e um dos objetivos do autor foi demonstrar em trabalhos subsequentes mais evidências que corroboram com esse modelo, como podemos observar na ilustração abaixo, publicada em 2011:

¹⁸ Como os autores mesmos colocam: “uma boa maneira de começar a considerar o comportamento geral do córtex cerebral é imaginar que a frente do cérebro está ‘olhando’ os sistemas sensoriais, a maioria dos quais está na parte de trás do cérebro” (Crick e Koch, 2003, pág. 120, tradução nossa); “um aspecto [perceptual] não pode tornar-se consciente, a menos que haja um nó essencial para ele [conexões com áreas frontais]. Para a consciência, pode haver outras condições necessárias, como projetar às áreas frontais do cérebro” (Ibidem, pág. 121, tradução nossa).

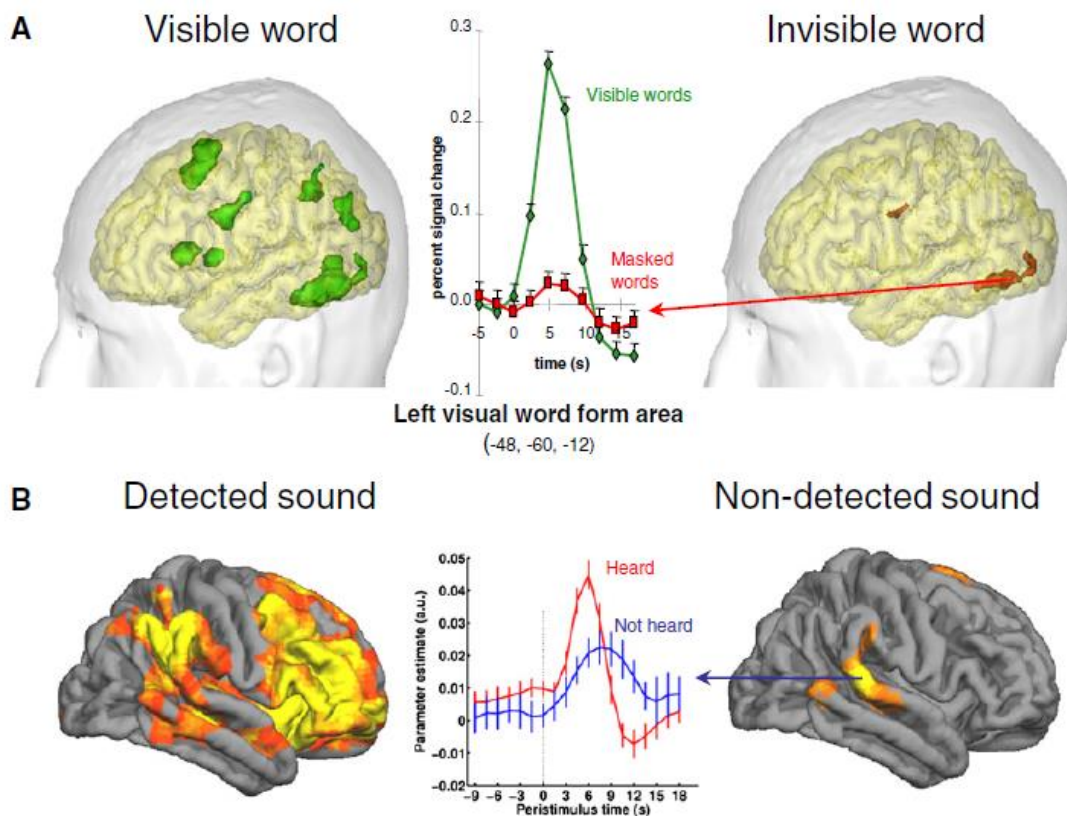


Fig. 6. Na ilustração, as partes A e B demonstram a diferença entre o processamento consciente (A) e inconsciente (B) de uma palavra; as partes C e D a diferença entre o processamento consciente (C) e inconsciente (D) de um estímulo sonoro. (Ilustração tirada de Dehaene e Changeux, 2011, pág. 203).

A imagem acima ilustra a distinção entre a atividade cortical de palavras percebidas e não percebidas conscientemente. Segundo o GNW, um estímulo quando experienciado conscientemente ativa conexões de longa distância e aumenta a atividade cortical [e Baars já tinha teorizado que o aumento da atividade é uma das evidências de atividade consciente, como dito anteriormente]. E novamente nós ressaltamos que a grande importância dada à atividade cortical quando se discute os dados empíricos sobre a atividade consciente, algo que veremos melhor quando apresentarmos as teorias de Koch e Tononi adiante.

Mais recentemente, Dehaene (2014, quarto capítulo) propôs quatro características (*signatures*) básicas da atividade consciente:

- Ampliação da atividade neural, principalmente nos córtices pré-frontais e parietais, em comparação à atividade inconsciente;
- Presença de P3 [um tipo de *potencial de eventos relacionados* (ERP) que indica responsividade cerebral para determinado estímulo sensorial medida através de eletro e magnetoencefalografia] após a apresentação de um estímulo. Não quando o estímulo é apresentado, mas sim quando ele é conscientemente experienciado, o que ocorre aproximadamente após 270 ms da apresentação do estímulo, possuindo picos de onda entre 350 e 500 ms (ou seja, possui um média de 300 ms, o que nomeia esse tipo de ERP: P300 ou P3). Para o autor, isso explica tanto ‘piscadas’ attentionais (*attentional blinks*) quanto a natureza serial da consciência, já que ela não é uma atividade contínua;
- Presença de ondas gama de alta intensidade após aproximadamente 300 ms da apresentação do estímulo. Como já discutido acima, Dehaene (2014) afirma que a hipótese das ondas gamas de 40 Hz de Crick e Koch (1990) foi uma ótima e forte hipótese, porém não é apenas a sua aparição que é importante, mas sim o fortalecimento das ondas, que ocorre quando o participante experiencia o estímulo conscientemente. Gaillard et al. (2009) demonstraram que uma palavra visível (experienciada conscientemente) está correlacionada com a sincronia de diferentes áreas corticais (como já dito acima) numa frequência acima de 30 Hz (ondas gama)¹⁹. Ver Fig. 6;
- Vasta sincronização neural. Gaillard e colegas (2009) aplicaram a análise de causalidade de Granger²⁰ e observaram que há uma causalidade bidirecional no

¹⁹ É interessante notar que ondas gamas são também detectadas durante sonhos (Llinás e Ribary, 1993).

²⁰ Seria possível argumentar que a análise de causalidade de Granger se refere apenas à correlação temporal, porém, nas neurociências, a correlação temporal é regularmente usada como um método para inferir correlações causais, por exemplo, se a atividade entre duas áreas cerebrais distintas demonstram padrões de ativação (estatisticamente significantes) onde uma é ativada após a ativação da outra se infere que a área que ativou primeiro controlou (ou causou) a ativação da outra.

cérebro em experimentos sobre consciência visual, existindo tanto uma sincronização causada por informações enviadas dos córtices visuais aos córtices pré-frontais, o que indicaria processamento consciente, quanto uma sincronização causada por informações enviadas dos córtices pré-frontais aos visuais, o que indicaria o controle *top-down* da atenção. E essa explosão causal, como os autores nomeiam, também ocorre por volta de 300 ms após a apresentação do estímulo. Logo, a consciência é também atestada quando há uma vasta sincronização neural, tanto de causalidade *bottom-up* quanto *top-down*.

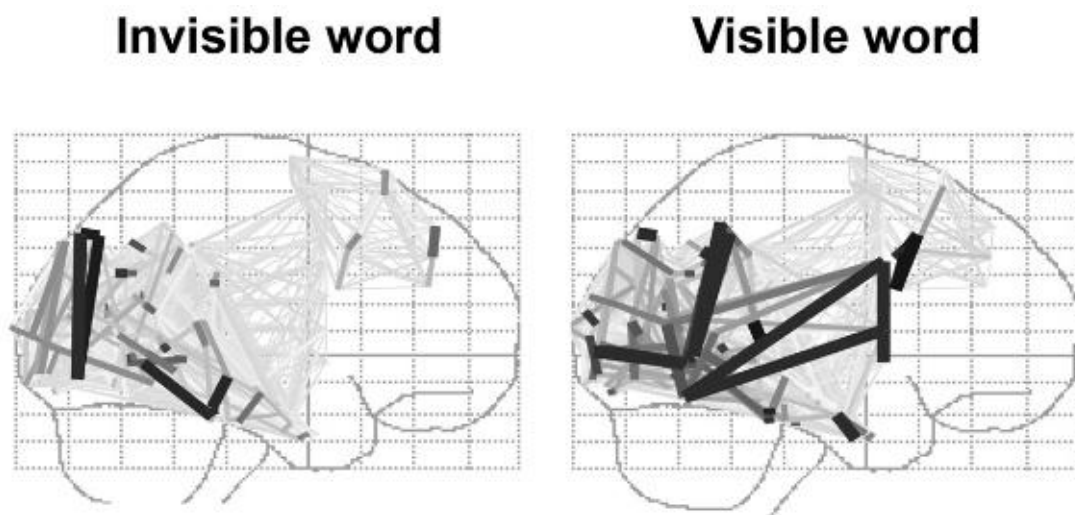


Fig. 7. Nesta ilustração podemos observar a diferença entre sincronizações com emissão de ondas gama fortes (linhas mais escuras) diante de palavras que são ditas como visíveis pelos participantes, ativando fortemente conexões de longas distâncias. (Ilustração tirada de Dehaene, 2014).

Em geral, um ponto importante a ser ressaltado aqui é a posição computacionalista da teoria de Baars e Dehaene. Ambas as teorias partem do pressuposto de que o cérebro é um sistema de codificação e decodificação, e que, de certo modo, possui programas. O viés computacionalista tanto da GWT quanto do GNW é explicitamente declarado tanto por Baars quanto por Dehaene. Por exemplo,

Baars utiliza o conceito de “biocomputador” para se referir ao cérebro no decorrer de seu livro de 1988 (como exemplo, ver Baars, 1988, pág. 45). Isso fica ainda mais explícito quando Baars e Franklin (2009) utilizam o modelo LIDA (*Learning Intelligent Distribution Agent*) para explicar a atividade consciente:

O progresso na consciência das máquinas dependerá da nossa compreensão da consciência nos cérebros humanos. Nós argumentamos que as funções da consciência biológica são implementadas de forma *bio-computacional*. Ou seja, os aspectos conscientes (assim como os não-conscientes) do pensamento, planejamento e percepção humanos são produzidos por *algoritmos adaptativos e biológicos*. Propomos que a consciência de máquina possa ser produzida por algoritmos adaptativos semelhantes aos que funcionam em uma máquina. (Baars e Franklin, 2009, pág. 9, tradução nossa, itálico nosso).

Além disso, Dehaene e colegas (2014) propõem um modelo para se entender a consciência de forma computacional. Analisemos a teoria de Baars sobre a atenção como exemplo.

Baars, juntamente com a maioria dos autores cognitivistas, trata a atenção como um controlador *top-down*, porém, nenhum deles propôs uma teoria ou modelo que explica nem por que a atenção é um controlador *top-down* nem como isso se dá, ou seja, eles não explicam a origem do controle atencional. Como pudemos observar acima, o modelo de atenção do Baars [que é o mais usado hoje em dia] parte do pressuposto que os estímulos competem entre si, onde aquele que se destaca ou vence atrai o *spotlight* atencional. Como explicado pelo próprio autor:

(...) Existe uma competição entre sistemas perceptivos para acesso ao espaço de trabalho global. Somente uma entrada [um input] pode ganhar, e é a única que garante o maior apoio de potenciais sistemas de recebimento, especialmente aqueles que estão prontos e “interessados” no sistema vencedor. “Ganhar” significa que um sistema ganha acesso ao tálamo e talvez à formação reticular, permitindo uma transmissão geral de alguns aspectos centrais do sistema vencedor - suas propriedades espaciais temporais, seu significado, sua relevância para os objetivos atuais, etc. Provavelmente alguns processadores receptores obtêm mais informações da mensagem global do que outras. Provavelmente existe um fluxo circulante de informações. Entre o sistema de entrada vencedor, o espaço de trabalho global e os processadores receptores, cada componente retornando para os outros,

de modo que há por algum tempo um ciclo auto-sustentável de sistemas ativados (...). (Baars, 1988, pág. 101, tradução nossa).

Como se pode observar, todo o modelo proposto pelo autor parte da competição entre estímulos, onde, havendo a competição entre dois estímulos, apenas um alcança o espaço de trabalho global e ganha o *spotlight* da atenção. Mas o que definitivamente controla esse *spotlight*? Os estímulos disputam igualmente? Há realmente uma disputa? A nosso ver, Baars, infelizmente, não oferece um modelo para compreender o funcionamento do *spotlight* da atenção, nem em termos computacionais (algoritmo). De qualquer forma, isso não limita a importância do modelo proposto por Baars em 1988.

E, seguindo esse mesmo modelo cognitivista-computacional, modelos que veem a consciência como um tipo de integração de informação também têm ganhado grande destaque na literatura, como o modelo de Giulio Tononi e Christof Koch.

Tonini e Koch

Giulio Tononi propôs, em 2004, o que ele chamou de Teoria da Informação Integrada (IIT) para explicar os fenômenos conscientes ao aplicar a teoria da informação à experiência perceptual, dando origem a novas ideias sobre o assunto, inclusive novos modelos matemáticos responsáveis por calcular quanta informação determinado sistema pode integrar.

Tononi (2004, 2008) sistematiza os princípios básicos da IIT argumentando que o modo correto de se teorizar sobre a consciência é a partir da experiência fenomênica. Segundo o autor (Tononi, 2008, pág. 217 e 218), quando temos uma experiência consciente ela não ocorre de modo compartimentado, por exemplo, quando olhamos para uma figura não temos como separar as formas das cores e experienciar cada uma dessas informações individualmente, ou seja, nosso cérebro integra diferentes tipos de informação numa mesma experiência consciente, e, já que isso é um modelo de

processamento de informação (integração de informação), o autor propõe que há como calcular o quanto de informação o sistema pode integrar.

Entretanto, Tononi (2004) afirma que o modelo de entropia da informação proposto por Shannon (1949), onde o mínimo de informação é representado por 1 bit, não é o modelo adequado para se calcular quanto de informação um sistema pode integrar, já que ele é unidirecional e não avalia as capacidades intrínsecas do sistema (Tononi, 2004, pág. 2 e 3). Por exemplo, segundo o autor, tanto um humano quanto um sensor fotossensível podem avaliar quando um monitor está aceso ou apagado, porém é óbvio que a experiência de um humano diante desses estímulos é completamente distinta do processamento de informação desse sensor, ainda mais se cada vez que a tela do computador estiver acesa ela mostrar imagens distintas. Nesse exemplo, o sensor apenas processa 1 bit de informação, ou seja, ou a tela está apagada ou acesa, já o humano integra um conjunto bem maior de informação, principalmente se, quando aceso, o monitor emitir diferentes estímulos visuais. Portanto, segundo Tononi (2004, 2008), não importa apenas o quanto de informação é transmitida, mas também como o sistema receptor responde à informação emitida, ou seja, é um modelo bidirecional.

Tononi (2004, 2008) percebe que o modo como o cérebro produz a consciência, ou seja, como integra informação (segundo o autor), também está relacionado à organização estrutural e funcional desse sistema. O autor reconhece que alguns tipos diferentes de lesões cerebrais resultam em diferentes perdas na experiência consciente, sem necessariamente causar perdas nos demais processos conscientes, como a perda da capacidade de ver cores, movimento, formas, etc., e que inclusive algumas dessas lesões impossibilitam essas experiências durante sonhos, enquanto que uma pessoa que perdeu a visão em estágios subsequentes do desenvolvimento ainda consegue sonhar com imagens visuais (Tononi, 2008, pág. 231). E isso também demonstra que podemos

analisar o processo de integração de informação por partes, a partir de como a informação é processada *dentro* do sistema e como as partes do sistema se comunicam e se influenciam *entre* si, e a partir disso podemos medir o quanto de informação o sistema pode integrar (Φ), o que é calculado através da entropia, estabelecimento de conjuntos e soma de matrizes²¹.

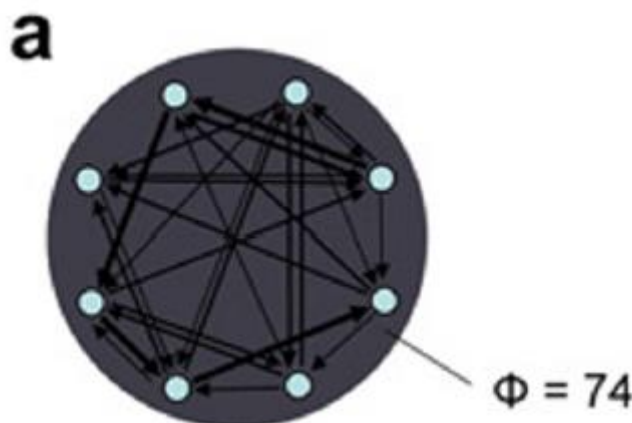


Fig. 8. A partir das relações entre os elementos de “a” podemos calcular o quanto de informação esse sistema pode integrar, no caso, 74 bits. Figura retirada de Tononi (2004, pág. 11).

O autor levanta algumas implicações para sua hipótese. A primeira é a capacidade de um sistema possuir consciência só depende da sua aptidão de integrar informação. Assim, segundo Tononi (2004, pág. 20), a consciência não depende de um *self*, linguagem, emoção ou corpo, mesmo admitindo que essas propriedades auxiliem no aprimoramento da capacidade consciente. Um exemplo disso, segundo o autor, é o sono REM, já que pessoas demonstram atividade consciente de modo isolado do ambiente externo e corporal (na maior parte das vezes) durante os sonhos.

Outra implicação, segundo Tononi (Idem, ibidem) é de que sistemas não-biológicos são capazes de ter atividade consciente, uma vez que o que realmente

²¹ É importante notar que o modelo apresentado por Tononi (2004, 2008) não é didático. O autor leva a metade do texto apenas para estabelecer a equação-base de seu modelo, envolvendo matemática aplicada. Caso se tenha curiosidade sobre como Tononi estabelece o cálculo de informação integrada, é melhor recorrer aos artigos originais em si. Como dito no início do capítulo, nosso objetivo é apenas fazer uma introdução (de certo modo inédita) a um debate muito mais extenso e complexo do que está apresentado aqui.

importa é a quantidade de informação que um sistema pode integrar e não simplesmente o arranjo organizacional entre os elementos do sistema (como os elementos do sistema estão dispostos e como eles se conectam), como representado na figura abaixo.

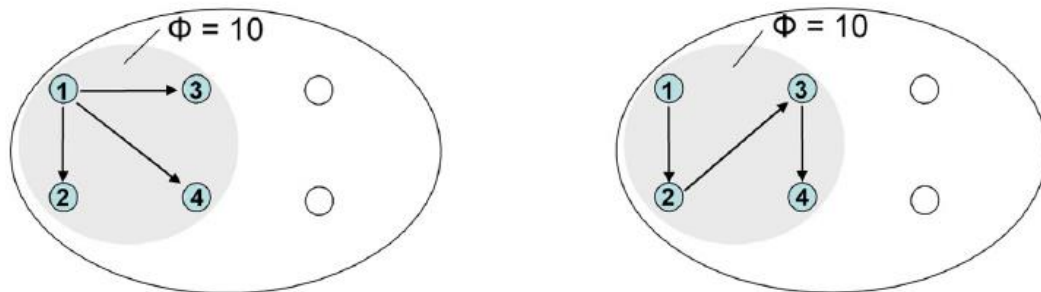


Fig. 9. Ambos os sistemas acima integram a mesma quantidade de informação mesmo possuindo interações distintas entre seus elementos. Ilustração retirada de Tononi (2004, pág. 8).

A teoria de Tononi teve um ótimo impacto acadêmico²², mas ela ganhou um apelo maior após o estabelecimento da parceria desse autor com Koch (que já era um pesquisador reconhecido e respeitado no assunto) na sistematização e divulgação da IIT, uma vez que Koch vem colaborando com a fundamentação neurobiológica dessa teoria. Mais recentemente, Tononi e Koch (2015) sistematizaram melhor as características do modelo utilizando-se de axiomas sobre a consciência fenomênica e postulados teóricos para explicar melhor a atividade consciente a partir de integração de informação (Tononi e Koch, 2010, pág. 7):

- Axioma 1: A consciência existe intrinsecamente, sendo sempre real e independente de observadores externos. Postulado 1: Desse modo, o sistema deve possuir poder de causa-efeito, existindo a partir de sua própria perspectiva intrínseca, além de ter poder de causa-efeito sobre si mesmo (interação entre seus elementos), o que os autores chamam de espaço de causa-efeito.

²² O artigo de 2004 e 2008 somam 927 e 771 citações respectivamente segundo o Google Scholar até a escrita desta dissertação.

- Axioma 2: A consciência é estruturada, já que cada experiência fenomênica é composta por diferentes elementos de alta-ordem. Postulado 2: O sistema também deve ser estruturado, onde seus elementos possuem diferentes relações uns com os outros (efeito causa-efeito sobre si mesmos).
- Axioma 3: Consciência é específica, onde cada experiência consciente é fenomenologicamente distinta das outras. Postulado 3: O sistema deve possuir um conjunto específico de repertórios de causa-efeito a partir de diferentes formas de relações entre seus elementos.
- Axioma 4: A consciência é unificada, toda experiência é formada pela integração de diferentes informações. Postulado 4: A estrutura de causa-efeito do sistema deve ser unificada.
- Axioma 5: A consciência é precisa, não é possível nem separar seu conteúdo em distintas partes nem desacelerar ou acelerar a experiência fenomênica. Postulado 5: As relações entre os elementos do sistema também devem ser precisas, não podendo haver nem mais nem menos relações de causa-efeito para gerar determinada experiência fenomênica.

De modo geral, a IIT aparenta ser uma teoria funcionalista (Tononi e Koch, 2015, pág. 15), mas ela não é tida assim pelos autores. O funcionalismo, como tido da filosofia da mente, além de considerar que um estado mental pode ser implementado por diferentes substratos, também defende que esse estado mental (software) não se reduz a sua implementação material (hardware) (Churchland, 1984/2013; Teixeira, 2000). Porém, esses autores se consideram reducionistas. Para eles, a experiência consciente não é uma consequência (causa) da interação de causa-efeito entre os componentes, mas sim a própria interação (a consciência não é tida como um fenômeno emergente). É

importante notar que são os próprios autores que não se consideram funcionalistas, o que fica visível nas seguintes passagens:

De acordo com muitas noções funcionalistas, se uma máquina reproduz nosso comportamento de entrada-saída em todas as circunstâncias, ter-se-ia que conceder a consciência para ela tanto quanto para nós. A IIT não podia discordar mais - nenhum teste de Turing (por exemplo, Samantha no filme de Hollywood *Ela*) pode ser um critério suficiente para a consciência, humana ou não. (Tononi e Koch, 2015, pág. 15, tradução nossa).

Finalmente, e a respeito de um computador cujo software simula em detalhes não apenas nosso comportamento, mas até mesmo a biofísica de neurônios, sinapses e assim por diante, da parte relevante do cérebro humano? Pode tal simulacro digital ser algum dia consciente? O funcionalismo novamente diria sim, ainda mais com força. Pois, neste caso, todos os papéis funcionais relevantes dentro do nosso cérebro, e não apenas o nosso comportamento de entrada-saída, teriam sido replicados fielmente. Por que não devemos conceder a este simulacro a mesma consciência que concedemos a um ser humano? De acordo com a IIT, no entanto, isso não poderia ser justificado, pelo simples motivo de que o cérebro é real, mas a simulação de um cérebro é virtual. Para o IIT, a consciência é uma propriedade fundamental de certos sistemas físicos, que exigem ter o verdadeiro poder causa-efeito, especificamente o poder de moldar o espaço de possíveis estados passados e futuros de uma maneira que seja maximamente irreduzível intrinsecamente. Do mesmo modo, massa é uma propriedade intrínseca de sistemas de partículas, uma propriedade que possui um poder de causalidade real, especificamente [o poder] de dobrar o espaço-tempo. Portanto, assim como uma simulação por computador de uma estrela gigante não vai dobrar espaço-tempo ao redor da máquina, uma simulação de nosso cérebro consciente não terá consciência (Idem, ibidem, tradução nossa).

Mesmo assim, os autores defendem que é possível existir diferentes qualidades de consciência, e diferentes níveis de integração de informação (diferentes Φ). No entanto, Tononi e Koch (2015) argumentam que os elementos básicos apresentados nos postulados acima devem estar presentes em todo sistema passível de possuir consciência, assim sendo, segundo os autores, o sistema *feed-forward* que temos hoje, como um computador comum, não é passível de possuir atividade consciente, já que ele não possui relações de causa-efeito entre seus componentes e é completamente dependente de sistemas externos, ou seja, não possui nenhum grau de adaptabilidade ($\Phi = 0$). Além disso, Tononi e Koch (2015) criticam a ideia de que o pansiquismo seja possível, uma vez que nesse “modelo” a consciência não teria nem começo nem fim – a

minha consciência não seria distinta da do leitor –, além de não explicar a relação entre *toda* a matéria e a mente (Ibidem, pág. 10).

Ainda mais recentemente, em 2016, Tononi, Koch e colegas publicaram uma *Perspective*²³ discutindo algumas formas de se investigar empiricamente os substratos físicos da consciência (Tononi et al., 2016), possuindo alguns pontos interessantes para ser discutido aqui. Nesse artigo os autores argumentam que o córtex posterior é a principal suplementação neural da consciência (bases corticais da consciência). Todo o artigo é embasado em como a organização *neuronal* do *córtex* poderia dar origem à consciência.

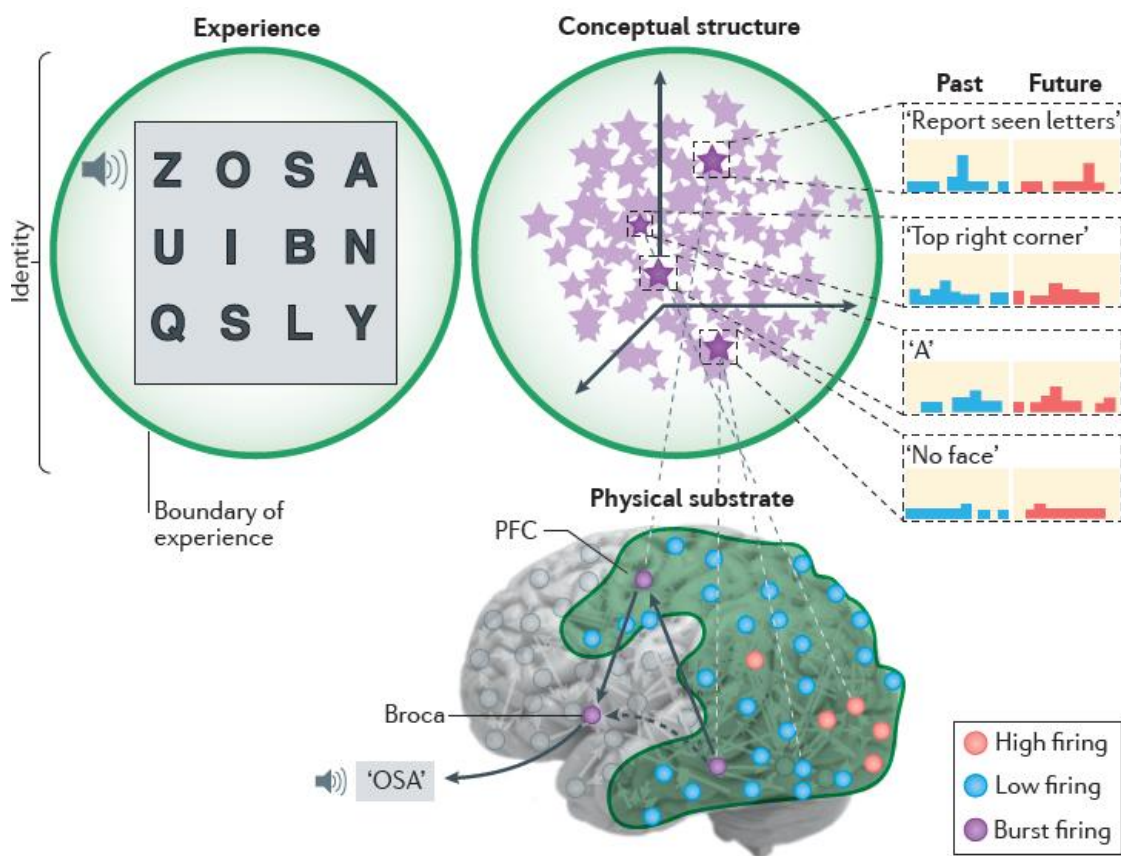


Fig. 10. Conteúdo fenomênico e acesso consciente. Essa imagem ilustra a atividade do córtex posterior diante do famoso teste de Sperling (1960), onde são apresentados brevemente ao participante arranjos de três a quatro letras do quadro acima. As estrelas representam o conteúdo fenomenal relatado pelo participante, expressões a partir dos gráficos de ativação cortical ao lado. O modelo cerebral representa os substratos físicos de tais experiências. Ilustração retirada de Tononi et al. (2016, pág. 458).

²³ “Perspective” (perspectiva) é uma forma de “review” (revisão de bibliografia) na qual os autores tratam o assunto abordado de modo mais pessoal, propondo nossas teorias ou modelos teóricos com base em dados empíricos já publicados. Esse artigo foi citado 83 vezes até o momento.

Pela imagem, podemos observar conexões entre os neurônios de diferentes regiões cerebrais [*physical substrate*] com diferentes níveis de ativação, como maior ativação numa área do córtex pré-frontal (PFC) e na área de Broca. Segundo os autores, o conteúdo consciente é dado por essa interconexão única entre esses neurônios, onde um interfere na atividade do outro. Desse modo, o conteúdo consciente se dá pela ativação única entre esses neurônios (resultando em diferentes níveis de ativação neuronal) e o acesso consciente se dá quando a informação é processada e transmitida entre eles. E, segundo a teoria, perdas nessas conexões devido a lesões, por exemplo, resultam em perdas no conteúdo e acesso consciente, ou seja, perdas na experiência consciente se dão devido a perdas na capacidade de integração do sistema.

Também recentemente, Koch, Tononi e colegas publicaram uma interessante *Review* (revisão) sobre o atual estágio das pesquisas dos correlatos neurais da consciência (Koch et al., 2016), oferecendo importantes informações para a presente dissertação²⁴. Uma delas é a “crítica” indireta às teorias subcorticais da consciência²⁵. Os autores apresentam um vasto número de dados através dos quais eles propõem que as áreas subcorticais não possuem causação direta (e em muitos casos, nem indireta) em relação à atividade consciente. O cerebelo, por exemplo, mesmo sendo uma área com vasto número de neurônios (quatro vezes mais neurônios que o córtex) e possuindo um grande número de conexões com outras áreas cerebrais, não possui nenhuma relação direta com a atividade consciente (Ibidem, pág. 309 e 310). Mesmo pacientes com agenesia²⁶ cerebelar demonstram atividade consciente praticamente intacta (Ibidem, pág. 310).

²⁴ Essa revisão já foi citada 139 vezes (em um ano).

²⁵ As seguintes informações são importantes para quando formos analisar as teorias de Damásio e Panksepp.

²⁶ Atrofia de um órgão ou tecido por parada do desenvolvimento na fase embrionária.

Já lesões no tronco cerebral²⁷ podem levar ao coma imediato, entretanto, pacientes com severas lesões corticais, mesmo possuindo o tronco cerebral intacto, permanecem em estado vegetativo (Idem, ibidem). O que, para os autores, demonstra que o tronco cerebral é necessário, porém não suficiente para sustentar a atividade consciente. Aliás, Koch e colaboradores demonstram uma postura da qual as áreas subcorticais, na verdade, são necessárias apenas para sustentar a atividade cortical no que diz respeito à produção da consciência, logo, lesões nessas áreas, na realidade, simplesmente impedem a capacidade do córtex de integrar informação (Idem, ibidem). Seguindo, os autores apresentam dados que demonstram que os gânglios basais também não possuem nenhum papel direto na produção da consciência, já que lesões bilaterais dessa área, mesmo dando origem ao mutismo, por exemplo, não impedem a responsividade a estímulos externos [os pacientes ainda demonstram estar conscientes desses estímulos]. E mesmo o claustró, que foi hipotetizado por Crick e Koch (2005) como uma importante área para a atividade consciente, aparenta não possuir poder causal [poder de causa-efeito discutido acima] sobre a atividade consciente, já que, como apontam os dados coletados por Damásio et al. (2013), pacientes com lesões bilaterais no claustró e áreas vizinhas (incluindo a ínsula) continuam demonstrando atividade consciente (Idem, ibidem).

E, por fim, os autores discutem o controverso papel do tálamo para a atividade consciente (Koch et al., 2016, pág. 310), argumentando que mesmo o tálamo sendo muito importante para a comunicação cortical, principalmente entre áreas corticais distantes (como entre o córtex visual e o pré-frontal), essa região não demonstra papel direto na produção da consciência, já que pacientes com atrofia talâmica não demonstram perdas na atividade consciente, e inconsciência induzida por anestésicos é

²⁷ Uma das regiões principais na teoria de Damásio, como veremos no segundo capítulo.

correlacionada com perdas na atividade cortical. Além disso, pacientes relatam alucinações hepinagógicas no início do sono quando o tálamo já está desativado (Magnin et al., 2010).

Já sobre as áreas corticais (Ibidem, pág. 310 e 311), os autores argumentam que os córtices primários, como o córtex visual primário, não possuem causação direta sobre a consciência. De acordo com eles, o córtex visual primário possui quase o mesmo poder de causa-efeito da retina, ou seja, não há evidências que corroborem seu papel causal na consciência visual. Mesmo para o córtex somatossensorial primário, que mais ou menos recentemente foi proposto como uma área importante para a experiência consciente (Meyer, 2011), ainda não há evidências definitivas que corroborem tal alegação.

Um ponto interessante desse artigo é quando os autores citam o *Opinion Article* de Safavi et al. (2014) que discute se o córtex pré-frontal é realmente responsável pela atividade consciente. Sem dúvida, há um grande número de trabalhos sobre o papel do córtex pré-frontal para a atividade consciente, como revisado por Dehaene e Changeux (2011), porém ainda não temos evidências que sistematizem a clamada função central dessa área cerebral para a consciência. Safavi e colegas discutem o experimento de Frässle et al., (2014) que, através de experimentos com pacientes com rivalidade binocular, demonstraram que quando os participantes experienciaram passivamente o resultado da rivalidade, sem a necessidade de realizar nenhum relato, a atividade no córtex pré-frontal ficou ausente²⁸, enquanto a atividade nos córtices visuais e parietais permaneceu, sugerindo, portanto, que o córtex pré-frontal está mais relacionado com a introspecção e produção de relatos do que com a percepção consciente em si. No entanto, Safavi et al. (2014) aconselham que nenhuma posição deve ser tomada de

²⁸ Baixo grau de atividade, estatisticamente ausente.

pronto no atual estágio do nosso entendimento sobre os correlatos neurais da consciência (NCC), uma vez que os experimentos de Frässle et al. não avaliaram a atividade de todo o córtex pré-frontal, além de ser necessário evidências adicionais para sabermos o real papel dessa região na atividade consciente, e é necessário fazer uma distinção entre áreas cerebrais relacionados com funções que são pré-requisitos para a experiência consciente e áreas que são realmente responsáveis por tal experiência. Ou seja, é importante discutir e pesquisar mais sobre a real importância das conexões frontoparietais na produção da consciência, que é atualmente quase um dogma na área. LeDoux (2015), por exemplo, afirma que essa posição é a mais corroborada hoje em dia, sendo um dos pilares da GWT e GNW, como visto acima.

Koch, Tononi e colegas, na verdade, argumentam em prol do que chamam de *posterior cortical hot zone* (Ibidem, pág. 311), propondo que o córtex posterior está mais relacionado com a atividade consciente do que o córtex pré-frontal. Segundos os autores, quando os participantes relatam sonhos durante o sono REM, por exemplo, a atividade do córtex pré-frontal é frequentemente mais baixa em relação à vigília. Os autores também afirmam que dados empíricos têm demonstrado que é difícil eliciar experiências sensoriais a partir de estimulações elétricas no córtex pré-frontal durante neurocirurgias, enquanto que esse tipo de experiência devido à estimulação nos córtices posterior é muito mais frequente, como percepção de faces e volição de mover um membro.

Além disso, dados neurológicos demonstram que pacientes com lobotomia frontal bilateral completa e com ressecções pré-frontais completas continuaram demonstrando atividade consciente (Idem, ibidem). Do mesmo modo, lobotomia e leucotomia pré-frontal, na qual essa área perde os inputs talâmicos, também não impedem a atividade consciente. Além disso, uma paciente com danos bilaterais

extensivos no córtex pré-frontal continuou demonstrando habilidade perceptiva e atividade consciente. Desse modo, os autores sugerem que os córtices pré-frontais estejam mais relacionadas com controle e execução de tarefas cognitivas do que com a experiência consciente em si (Koch et al., 2016, pág. 311).

Já sobre dados neurofisiológicos, os autores problematizam alguns “dogmas” das pesquisas sobre NCC, como ondas gama e P3²⁹. Koch e colegas (Ibidem, pág. 313) argumentam que, desde a famosa publicação de Gray et al. (1989), ondas gama (30-70 Hz) vêm sendo tidas como correlatos de atividade consciente por inúmeros autores, porém a maioria dos trabalhos sobre atividade consciente e ondas gamas não levam em consideração a distinção entre consciência e atenção seletiva, uma vez que há registro de ondas gama mesmo quando o participante alega não ter visto o estímulo, e elas também estão presente no início do sono NREM e durante efeitos anestésicos, o que sugere que ondas gama também estão presente quando não há atividade consciente. No entanto, Dehaene (2014), como apresentado acima, admite esses dados, mas argumenta que o que realmente importa é o fortalecimento dessas ondas quando o participante relata experiência consciente, ou seja, o que importa é o fortalecimento desse tipo de onda e não apenas sua presença, algo que, portanto, deve ser mais investigado.

Outro tipo de dado eletrofisiológico, comentado acima, é uma ERP (potencial de eventos relacionados) frontoparietal que surge aproximadamente após 300 ms da apresentação do estímulo, conhecida como P300 ou P3. Porém, os autores argumentam que esse tipo de ERP não está presente quando os participantes não desempenham nenhuma tarefa, mesmo estando conscientes do estímulo, ao mesmo tempo em que estímulos não percebidos conscientemente podem disparar esse tipo de ERP. Além disso, a presença de P3 não difere se um paciente está em coma vegetativo ou

²⁹ Discutidas acima quando analisamos os trabalhos de Dehaene.

minimamente consciente, além de não ser percebida em pacientes com determinadas lesões cerebrais, mesmo quando eles continuam conscientes (Koch et al., 2016, pág. 313).

Ou seja, Koch, Tononi e colegas argumentam contra alguns “dogmas” das pesquisas sobre NCC em prol de sua teoria, a IIT. Mas, os autores continuam defendendo que a base neural da consciência é estritamente cortical, sendo as áreas subcorticais apenas o suporte (responsáveis por atividades necessárias à consciência) da atividade cortical consciente, algo que devemos ter em mente quando formos analisar as teorias de Damásio no segundo capítulo.

De qualquer forma, as teorias cognitivistas da consciência são também embasadas em outros tipos de modelos filosóficos cognitivistas, como o modelo de LeDoux e Brown (2017), fundamentado na teoria de alta-ordem da consciência (HOT), que, por ser uma nova teoria sobre consciência emocional, tem ganhado grande visibilidade na área da neurobiologia das emoções e consciência emocional, devendo, então, ser também brevemente analisada na presente dissertação.

LeDoux

Antes de 2012, LeDoux também foi um autor que defendia que as emoções eram processos inconscientes e diretamente vinculados aos circuitos de sobrevivência do sistema nervoso, como pode ser visto em seu importante livro, *The emotional brain*, de 1996. Porém, em 2012, LeDoux publicou uma importante *Perspective* na revista *Neuron*³⁰, pela qual o autor reformulou a visão tradicional de emoção. LeDoux trabalhou essa reformulação (nova tese) em artigos posteriores (LeDoux, 2014; LeDoux e Pine, 2016; LeDoux e Brown, 2017; LeDoux, 2017) e no seu livro de 2015, *Anxious*.

³⁰ Citada 732 vezes até o momento.

Mas o real motivo para analisarmos as ideias de LeDoux nessa seção é que, em 2017, LeDoux e Brown³¹ propuseram um novo modelo de consciência emocional. Logo, uma breve análise das ideias de LeDoux aqui fez-se muito necessária.

Como será analisado na seção “o paradigma afetivista” abaixo, é comum na literatura sobre emoções vincular os estados emocionais com os circuitos de sobrevivência inatos do organismo, de modo que muitos autores, como Panksepp, Frijda, Ekman e Damásio (este último veremos melhor posteriormente), defendem que determinados estados emocionais são inatos (comumente chamados de emoções básicas, mesmo que os autores discordem sobre quais especificamente são essas emoções básicas). E, como veremos abaixo, Panksepp (2000), com base nessa ideia central, chega a propor que as emoções básicas são *espécies naturais*, de modo que haja um circuito subcortical único dedicado às emoções, sendo este compartilhado por inúmeras espécies.

Porém, no artigo de 2012, LeDoux propõe uma distinção epistemológica entre emoção e circuitos de sobrevivência. Segundo o autor, as reações inconscientes que ocorrem no sistema nervoso antes que tenhamos uma compreensão consciente do estímulo causador deve ser classificada como *ativação dos circuitos de sobrevivência*. Essa distinção tem inúmeras consequências para a nossa discussão aqui. A mais importante delas é que processos tidos como “afetivos” pela visão tradicional de emoção, como valência, motivação e excitabilidade, são, na verdade, mecanismos inconscientes associados com a ativação dos circuitos de sobrevivência, sendo apenas *indiretamente* vinculados com os estados emocionais. O autor baseia esse posicionamento no fato de que os mecanismos neurais responsáveis por esses processos psicofisiológicos são ativados antes que tenhamos qualquer tipo de experiência

³¹ É interessante notar que Richard Brown é, na verdade, um filósofo.

consciente sobre o estímulo. Por exemplo, tanto valência quanto motivação estão correlacionadas com a ativação do núcleo accumbens, que é uma estrutura subcortical da qual a ativação não depende que o organismo tenha experiência consciente do estímulo, e a ativação dessa área não necessariamente dá origem a uma experiência emocional consciente (Winkielman et al., 2005). E excitabilidade é controlada pelo sistema neuromodulatório central e pelo sistema hormonal periférico (LeDoux, 2012, pág. 662), que são responsáveis por inúmeros processos relacionados com a sobrevivência, que nem sempre dão origem a algum tipo de experiência consciente. Ou seja, LeDoux quebra com a tradição de atribuir uma natureza afetiva ou emocional a processos subcorticais *inconscientes* relacionados diretamente com a sobrevivência do organismo.

Nesse artigo, e no livro de 2015, LeDoux dialoga diretamente com o modelo de espaço global para se referir à consciência. Como argumenta o próprio autor: “a ideia proposta aqui é que os sentimentos conscientes resultam quando os estados organísmicos globais são representados no *espaço de trabalho cognitivo*” (LeDoux, 2012, pág. 665, tradução nossa, *italico nosso*). Desse modo, LeDoux defende um modelo puramente cognitivista de emoção, a partir do qual a ativação dos circuitos de sobrevivência e outros processos inconscientes não são afetivos ou emocionais por si sós. A criação do estado emocional só se dá quando esses processos são vinculados a outros processos cognitivos, como memória explícita e linguagem, no espaço de trabalho global, como ilustrado na imagem abaixo:

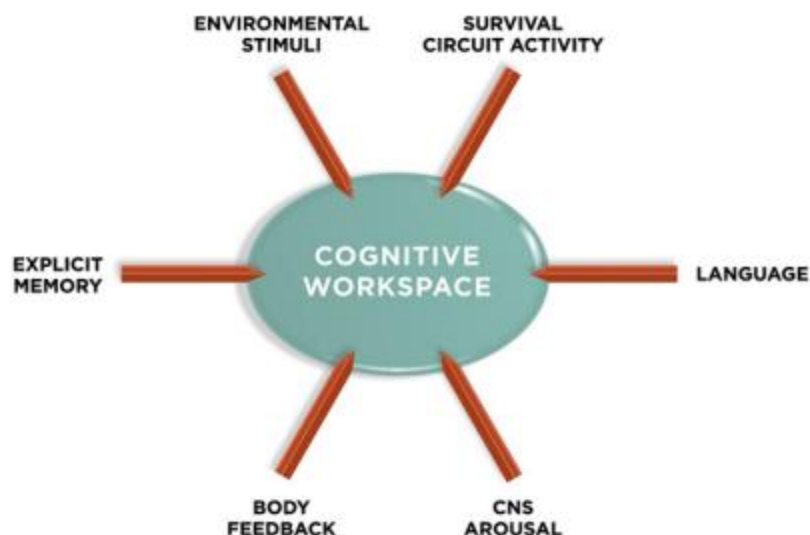


Fig. 11. Ingredientes do sentimento no espaço de trabalho global. Imagem tirada de LeDoux (2012, pág. 666).

Porém, LeDoux não segue adiante com esse modelo cognitivista especificamente, aderindo, em 2017, à *teoria de representação de alta-ordem de uma representação* (teoria HOROR) de Richard Brown (2015), que, de qualquer forma, continua sendo uma teoria cognitivista. Logo, antes de discutirmos o modelo de consciência emocional de LeDoux e Brown (2017), temos de analisar os fundamentos principais dessa teoria cognitivista da consciência.

A teoria HOROR, assim como qualquer teoria de alta-ordem da consciência (HOT), é baseada no modelo de re-representação de uma representação de primeira-ordem não-consciente. Seguindo o modelo HOT de consciência, a experiência visual, por exemplo, só se dá quando a representação visual realizada pelo córtex visual (representação de primeira-ordem) é re-representada por mecanismos de alta-ordem, como os mecanismos pré-frontais. Porém, a teoria HOROR de Brown (2015) admite que processos não-sensoriais podem também ser re-representados, como re-representação de uma representação de alta-ordem, dando origem a uma segunda representação de alta-ordem: uma representação de pensamento (pensamento sobre uma

representação de alta-ordem), que, segundo o autor, caracteriza a introspecção. É importante notar que, para qualquer tipo de teoria de alta-ordem, o *self* possui um papel fundamental. Não haveria como ter uma representação de pensamento sem *self*, para haver consciência, a representação de alta-ordem sempre deve ser direcionada ao *self* (Rosenthal, 2005; LeDoux, 2015; Brown, 2015; LeDoux e Brown, 2017). Mas, LeDoux e Brown (2017) corretamente ressaltam que isso não significa que a teoria HOROR (ou qualquer outra HOT) se refere apenas a um modelo de autoconsciência. Os autores ressaltam a diferença entre “estar consciente de um estado” (o que também envolve *self*) e “estar consciente que eu estou consciente de um estado mental” (o que envolve um *self* que tem consciência de si mesmo) (LeDoux e Brown, 2017, pág. E2020). Esse segundo tipo de consciência (autoconsciência) é chamado de *self-HOROR* no modelo de LeDoux e Brown (2017).

De qualquer modo, o objetivo central de LeDoux e Brown é, na verdade, utilizar o modelo da teoria HOROR de Brown (2015) para propor um novo modelo de consciência emocional, o que os autores chamam de *teoria de alta-ordem da consciência emocional* (HOTEC). Como visto acima, um estado mental só se torna consciente quando este é re-representado por processos de alta-ordem. Os autores partem do ponto que, já que a atividade dos córtices pré-frontais está correlacionada com acesso consciente (Frith et al., 1999; Dehaene et al., 2006; Dehaene e Changeux, 2011; Fleming et al., 2012; Dehaene, 2014), é provável que são essas áreas as responsáveis por re-representar as representações de primeira-ordem, além de outras representações de alta-ordem (teoria HOROR). O conjunto de áreas responsáveis por realizar tais re-representações é chamado pelos autores de *redes gerais de cognição* (GNC), constituídas pelos córtices pré-frontais juntamente com a ínsula e o córtex parietal posterior (LeDoux e Brown, 2017, pág. E2018).

De fato, há um conjunto de evidências que corroboram a posição teórica de LeDoux e Brown, já que a ativação dos circuitos de defesa, como ativação da amígdala, não é o suficiente para dar origem ao sentimento de medo (Morris et al., 1999; Williams et al., 2006; Bornemann et al., 2012), o que também ocorre com pacientes com visão cega (Morris et al., 2001; Tamietto e de Gelder, 2010). Além disso, danos bilaterais à amígdala não é o suficiente para interromper estados conscientes de medo e ansiedade (Anderson e Phelps, 2002; Feinstein et al., 2013). Logo, segundo LeDoux e Brown (2017), é provável que a ativação dos circuitos de sobrevivência apenas forneça representações (não-conscientes) de primeira-ordem que podem ou não participar na construção do estado emocional de medo pelas GNC.

Nesse artigo, LeDoux continua defendendo que o estado emocional se dá apenas a partir da integração de um conjunto de processos cognitivos e fisiológicos na memória de trabalho, tanto que, segundo LeDoux e Brown, todo processo emocional (em humanos, pelo menos) são baseados em *esquemas emocionais do self*, que são um conjunto de informações sobre uma determinada emoção, por exemplo, memórias autobiográficas sobre o que significa o sentimento de medo, o que envolve memória, linguagem, mecanismos de rotulagem (qual emoção estou sentindo), etc. (LeDoux e Brown, 2017, pág. E2020). Logo, de acordo com essa teoria, não haveria processos emocionais sem a integração de processos fisiológicos e cognitivos. Portanto, ao contrário de Panksepp (como veremos a seguir), por exemplo, LeDoux e Brown não defendem que possam existir processos afetivos ou emocionais intrínsecos a mecanismos subcorticais. Para esses autores, assim como para Barrett (como veremos no terceiro capítulo), todo processo emocional é *cognitivamente construído*. A imagem abaixo ilustra os pontos gerais da teoria HOTEK de LeDoux e Brown (2017):

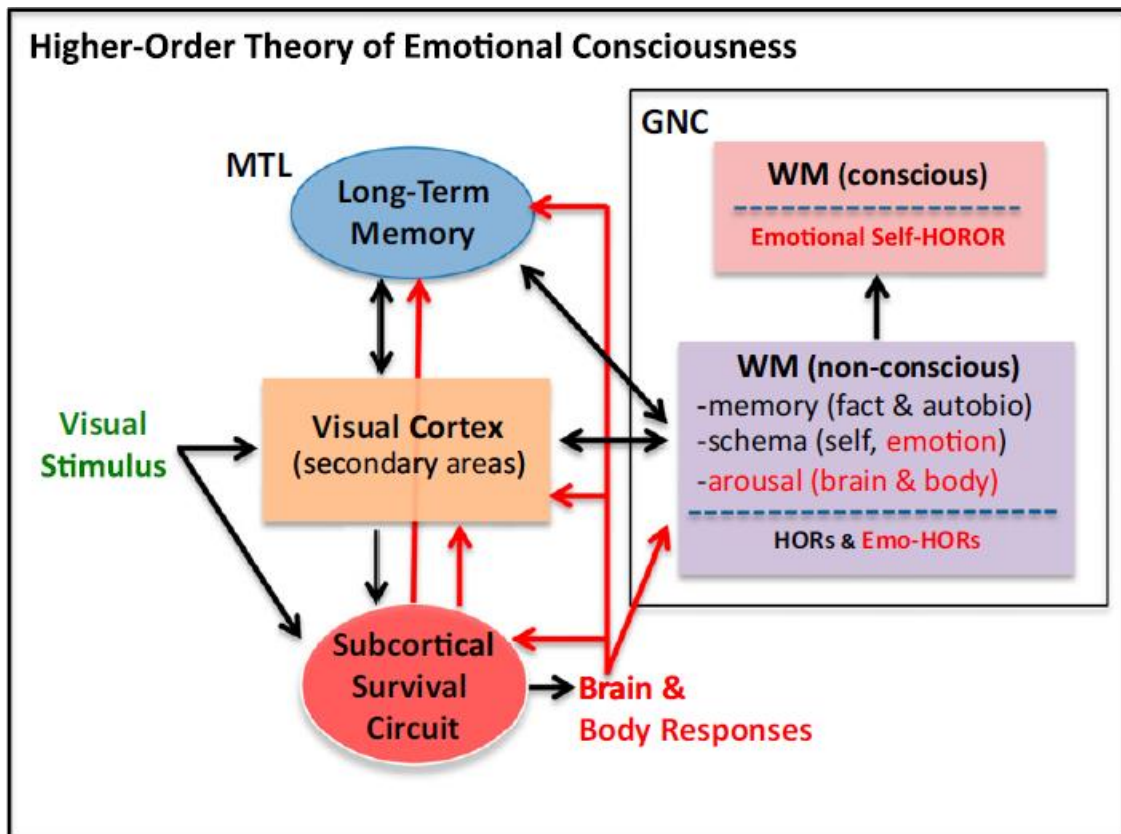


Fig. 10. Modelo HOTEK de construção dos estados emocionais. MTL significa lobo temporal medial. WM significa memória de trabalho. Figura retirada de LeDoux e Brown (2017, pág. E2021).

A consciência cognitiva

Mesmo que as teorias apresentadas acima sejam distintas, todas elas compartilham o mesmo paradigma, de modo que podemos levantar algumas ideias centrais inerentes a todas elas.

Como pudemos ver, todas as teóricas cognitivistas são, acima de tudo, corticalistas, e isso não ocorre por acaso. Como mencionado no início dessa seção, o estudo realizado por Penfield e Boldrey (1937) demonstrou que as representações sensoriais são vastamente processadas por áreas corticais, e, já que a percepção é caracterizada pela integração de dados sensoriais (Pomerantz, 2003), não seria possível, seguindo essa visão, haver consciência perceptual sem processamento cortical. Além disso, como muito bem analisado por Dehaene (2014) e LeDoux e Brown (2017),

inúmeros processos cognitivos necessários à consciência, como memória de trabalho, metacognição e atenção, vêm sendo correlacionados com atividades corticais, principalmente no córtex pré-frontal. Desse modo, segundo esse modelo, mesmo que as áreas subcorticais providenciem importantes representações para a construção da consciência (como atividade do tronco cerebral superior, que inclui o tálamo), não haveria consciência (perceptual e emocional) sem atividade cortical (Dehaene, 2014; LeDoux, 2015; Koch et al., 2016; LeDoux e Brown, 2017).

Além disso, três das teorias apresentadas acima seguem uma visão computacionalista de consciência, no caso, a GWT, a GNW e a IIT (mesmo que isso não seja compartilhado por todos os modelos cognitivistas, como o modelo de Crick e LeDoux). Isso está embasado na ideia de que o cérebro é responsável por codificar, decodificar e integrar informação interna e externa. A nosso ver, é possível que essa visão geral tenha dado origem ao “cerebrocentrismo” do paradigma cognitivista, já que a maioria dessas teorias, mesmo admitindo a importância de feedback corporal (como o modelo de LeDoux), postula que os *neurônios* corticais são os responsáveis por construir a consciência (segundo o modelo de LeDoux, por exemplo, o feedback corporal é importante, mas não fundamental). Essa posição chega a dar origem a teorias um tanto quanto radicais, como a IIT de Tononi e Koch, como visto acima, já que os autores postulam que a atividade consciente não necessita de corpo, emoções e *self*.

Mas, em contraposição ao paradigma cognitivista, há o paradigma afetivista, que admite que a consciência afetiva é inerente à atividade subcortical relacionada com os mecanismos de sobrevivência. Segundo o paradigma afetivista, em termos gerais, a consciência é majoritariamente subcortical e não-computacional.

O paradigma afetivista da consciência

Antes de analisarmos a teoria afetivista de Panksepp em si, há alguns conceitos que devem ser abordados aqui: valência, motivação e excitabilidade, já que as teorias sobre o afeto são baseadas nesses processos psicofisiológicos (Frijda, 1986). Valência está relacionada à avaliação subjetiva da situação na qual o organismo se encontra, podendo ser tanto positiva quanto negativa. De certo modo, valência pode ser classificada como uma “atratividade” (positividade) ou “aversão” (negatividade) ligada à própria natureza do estímulo do ponto de vista de cada *organismo* (Frijda, 1986, pág. 207). Desse modo, algo que pode ter uma valência positiva para determinado organismo (como o estímulo “carniça” para um urubu) pode ter valência negativa para outro organismo (como o estímulo “carniça” para um ser humano saudável). Logo, inúmeras teorias emocionais (tanto afetivas, como as teorias de Panksepp, quanto cognitivas, como as teorias de Barrett, 2006a) tomam valência como um conceito importante para classificar a natureza das reações emocionais, por exemplo, tomando a emoção tristeza como uma reação de valência negativa e a emoção alegria como uma reação de valência positiva³².

O conceito de valência está intrinsicamente ligado ao conceito de motivação. Segundo Robbins e Everitt (1996), motivação pode ser entendida como antecipação do reforço, seja ele positivo³³ ou negativo³⁴, onde o organismo busca ou evita determinado estímulo devido a seu histórico de reforçamento. De certo modo, podemos dizer que motivação positiva está ligada à positividade da valência, ou seja, um estímulo com

³² De qualquer modo, é importante ressaltar que esses autores defendem modelos de emoção significativamente distintos, algo que será melhor analisado no terceiro capítulo. Panksepp (1998, 2000, 2005), por exemplo, defende que as emoções básicas [e aqui é importante ressaltar que as emoções básicas defendidas por Panksepp são distintas das defendidas por Ekman, 1992] são inatas e pertencem a todos os organismos que possuem sistema nervoso (esse autor defende que emoções são *espécies naturais*). Já Barrett (2006, 2017) defende a “teoria das emoções construídas”, segundo a qual não há emoções inatas. Segundo essa teoria, todas as reações emocionais são construídas pelo contexto do qual o organismo está inserido através da linguagem. A autora inclusive chega a comparar o conceito de emoção ao conceito de dinheiro, sendo ambos construídos socialmente (Barrett, 2017, pág. 13). Uma breve análise das ideias de Barrett será realizada no terceiro capítulo desta dissertação.

³³ Que aumenta a probabilidade da resposta devido à presença (positividade) da recompensa.

³⁴ Que aumenta a probabilidade da resposta em relação à ausência (negatividade) o estímulo.

valência positiva desencadeia uma motivação de aproximação, enquanto um estímulo com valência negativa desencadeia uma motivação de evasão/fuga.

E excitabilidade está ligada a intensidade da situação em si. Por exemplo, reações de alegria intensa e excitação sexual estão associadas com grande excitabilidade nervosa, enquanto tristeza e luto estão associados com baixa excitabilidade nervosa. Excitabilidade nervosa é geralmente associada com a resposta de fuga ou luta, controlada, majoritariamente pelo sistema nervoso autônomo (Panksepp, 1998, pág. 44).

Como podemos ver, esses processos acima estão vinculados com a sobrevivência do organismo (Frijda, 1986). A interação dos organismos com o ambiente depende da avaliação da situação (valência), da aproximação a estímulos positivos e evasão de estímulos negativos (motivação) e cada uma dessas situações modula o contexto fisiológico do organismo (excitabilidade). E isso é fundamental para entendermos os pontos centrais do paradigma afetivista de consciência e emoção, que possui, como principal representante, Jaak Panksepp.

Panksepp

Em termos gerais, podemos dizer com grande convicção que Panksepp é, hoje em dia, o principal autor do paradigma afetivista da emoção³⁵. Panksepp desenvolveu de modo bem denso e completo um método de observar e analisar o comportamento afetivo a partir de modelos animais, amplamente discutidos em seu livro *Affective Neuroscience* (1998), que é uma leitura fundamental para todos os interessados na discussão sobre emoção e consciência animal. De modo geral, as ideias de Panksepp podem ser divididas em duas grandes fases: os trabalhos de Panksepp antes de 2009, ligados a uma visão mais conservadora de afeto, e os trabalhos após 2010, que começam

³⁵ Seu livro *Affective Neuroscience* (1998) foi citado 7303 vezes segundo o Google Scholar até a escrita dessa dissertação.

a dialogar mais com a neurociência cognitiva, *dado as devidas restrições*, já que o autor continuou mantendo alguns axiomas centrais de sua teoria, como veremos posteriormente.

Até por volta de 2009, as obras de Panksepp se aproximavam muito mais do modelo afetivista proposto por Zajonc (1984), que separava os processos cognitivos dos processos afetivos, de modo que seria possível ocorrer processos afetivos sem a necessidade de processos cognitivos. Por exemplo, em resposta ao artigo de Gray (1990), que propõe que processos emocionais não são independentes de processos cognitivos, Panksepp (1990) defende que processos afetivos podem ser separados dos processos cognitivos. Nesse artigo, Panksepp declara que processos afetivos possuem, por eles mesmos, o poder de eliciar comportamento, que não é necessário ter uma compreensão clara dos processos cognitivos para compreender os processos afetivos (Ibidem, pág. 289) e que os processos emocionais são processos irreduzíveis (provavelmente no sentido de não serem redutíveis a processos cognitivos) (Ibidem, pág. 291)³⁶.

Essa posição de Panksepp está completamente de acordo com a discussão acima sobre a relação entre valência, motivação, excitabilidade e sobrevivência. Panksepp (1990, 1998, 2000, 2005) defende que os processos emocionais são herdados pela evolução, de modo que o autor considera que as emoções básicas são *espécies naturais* (Panksepp, 2000). Se nós consideramos, por exemplo, que os processos emocionais são herdados através do processo evolutivo, isso está certamente baseado na ideia de que emoção é um fenômeno relacionado à sobrevivência do organismo (se emoção foi selecionada, é porque certamente esse padrão de comportamento aumentou a taxa de

³⁶ É importante notar que, nesse debate, o artigo de Gray alcançou um impacto significativamente maior que o artigo de Panksepp sobre a mesma questão. O artigo de Gray (1990) foi citado 1377 vezes segundo o Google Scholar até a escrita desse texto, enquanto o artigo de Panksepp foi citado 41 vezes no mesmo período de tempo. Isso pode ter ocorrido devido à força do paradigma cognitivista, como comentado acima.

sobrevivência do organismo). Logo, não é coincidência que Panksepp utilizou os conceitos de valência, motivação e excitabilidade ao longo do *Affective Neuroscience* (1998). Na verdade, a relação entre emoção e sobrevivência foi adotada por inúmeros autores proeminentes na área (Darwin, 1872; Smith e Lazarus, 1990; Ekman, 1992; Damasio, 1994, 1999, 2010; LeDoux, 1996; Panksepp, 1998; Keltner e Gross, 1999; Damasio e Carvalho, 2013; para revisão nesse tema veja LeDoux, 2012). Por exemplo, há autores que defendem que o sistema nervoso autônomo é responsável por desencadear processos emocionais (Ekman, Levenson e Friesen, 1983; Panksepp, 1998; Kreibig, 2010; Levenson, 2014). Logo, de um modo ou de outro, o paradigma afetivista está intrinsicamente vinculado à visão de que processos emocionais são processos diretamente ligados à sobrevivência do organismo.

Levando essa posição a sério, Panksepp (2000) argumenta que existem circuitos (subcorticais) inatos dedicados aos processos emocionais, propondo que emoções são espécies naturais (*natural kinds*). Panksepp chega a essa conclusão a partir da observação direta do comportamento “afetivo” de outras espécies, como outros mamíferos e alguns moluscos, como polvo. Entretanto, é importante ter em mente que as emoções básicas das quais Panksepp (1998, 2000, 2005) se refere não são as mesmas emoções propostas por Ekman (1992). Isto é, Panksepp não se refere à tristeza, alegria, raiva, medo, amor, gosto e desgosto (Ekman, 1992). Panksepp, na verdade, traça padrões de comportamento como características básicas das reações emocionais, caracterizados pela forma de como esses animais se comportam, ao invés de se basear nos modelos humanos de experiência emocional. Desse modo, Panksepp (1998, 2000, 2005) classifica sete comportamentos emocionais básicos: busca, medo, raiva, luxúria, cuidado (no sentido de proteger outro indivíduo), pânico e brincadeira. Segundo o autor, esses comportamentos podem ser observados em inúmeras espécies, sugerindo,

portanto, que essas espécies podem sim ter consciência afetiva. Logo, para o autor, é o comportamento que indica se o organismo possui ou não consciência afetiva, e é importante ter essa informação em mente.

A visão de que os processos emocionais estão intrinsecamente vinculados aos mecanismos de sobrevivência também dá origem à posição de que os processos emocionais podem ser instanciados por mecanismos subcorticais. Como se sabe, o neocórtex é a região mais recente de todas as estruturas cerebrais, e essa estrutura corresponde a 90% do total da massa cortical humana (em humanos, o neocórtex corresponde a 76% do cérebro) (Noback et al., 2005). Portanto, seguindo a posição de que emoções são processos neurobiológicos herdados através do processo evolutivo, é muito improvável que, anteriormente, esses processos tenham sido instanciados por mecanismos neocorticais, sendo, desse modo, provavelmente instanciados por processos subcorticais (Panksepp, 1998). De qualquer forma, não haveria como defender que as emoções básicas são espécies naturais ao mesmo tempo em que se defende uma visão (neo)cortical de emoção (isto é, se as emoções básicas são espécies naturais, elas estão presentes em inúmeros organismos, tanto naqueles que possuem córtex quanto naqueles que não o possuem).

Porém, uma dúvida permanece: como que Panksepp argumenta que consciência afetiva podia ser instanciada por processos subcorticais sem a necessidade de processos cognitivos? Atualmente, já está estabelecido na literatura que as áreas corticais são fundamentais para inúmeros processos cognitivos, incluindo visão, audição, comunicação, interocepção, etc., ao menos em humanos e outros mamíferos (Miller e Cohen, 2001; Elston, 2003; Pascual-Leone et al., 2005). A posição teórica que processos afetivos não dependem de processos cognitivos está compatível com a visão de que mecanismos subcorticais podem, por si mesmos, instanciar processos afetivos.

Mas como seria possível argumentar que consciência, afetiva ou de qualquer outro tipo, pode ser instanciada por processos subcorticais sem a necessidade dos processos cognitivos corticais³⁷?

Panksepp (1990, 1998, 2005) deixa claro que ele não desconsidera a importância da interface afetivo/cognitivo para se compreender os processos afetivos, por exemplo, o autor considera que processos cognitivos podem inibir processos afetivos, e que processos afetivos contribuem em processos cognitivos, como tomada de decisão (Panksepp, 1990, pág. 290; 1998, pág. 150; 2005, pág. 62). No entanto, até 2009, Panksepp continuava considerando que processos emocionais podem ocorrer sem a necessidade de processos cognitivos. Para defender a autonomia da consciência afetiva em relação a processos cognitivos, Panksepp (2005, pág. 32) faz uso do monismo de duplo aspecto de Spinoza. Não é necessário que o leitor tenha um vasto conhecimento sobre esse modelo para entender a posição de Panksepp. Para Panksepp (2005), o afeto é em si um dos aspectos dos processos neurais, ou seja, “(...) os afetos emocionais podem estar completamente incorporados nas atividades ampliadas dos sistemas operacionais do cérebro que orquestram certos padrões de ação instintiva” (Panksepp, 2005, pág. 38, tradução nossa). Desse modo, Panksepp propõe que áreas subcorticais responsáveis por respostas instintivas e automáticas são também responsáveis por instanciar processos afetivos, já que o afeto é um aspecto inerente à atividade neural dessas áreas cerebrais³⁸. Portanto, através desse posicionamento filosófico, Panksepp responde por que mecanismos subcorticais podem realizar atividade consciente sem a necessidade de avaliação cognitiva cortical.

³⁷ Aqui cabe ressaltar que, mesmo que o tálamo venha sendo correlacionado com inúmeros processos cognitivos (Saalman e Kastner, 2015), isso não significa que podemos menosprezar a importância dos processos cognitivos corticais, como analisado na seção anterior.

³⁸ É importante ressaltar que o próprio Panksepp não discute de modo aprofundado a teoria do monismo de duplo aspecto nesse artigo.

E cabe notar que, para Panksepp, o estado afetivo é sempre um estado *consciente*. Para o autor, o afeto é intrinsecamente ligado ao aspecto qualitativo da mente. Como o autor mesmo coloca:

(...) *afeto* é o componente experiencial-sentimental subjetivo que é muito difícil de descrever verbalmente, mas há uma variedade de afetos distintos, alguns vinculados de forma mais criticamente aos eventos corporais (unidades homeostáticas como fome e sede), outros a estímulos externos (gosto, toque, etc.). (Panksepp, 2005, pág. 32, tradução nossa, itálico do autor).

Desse modo, podemos dizer que Panksepp vincula o conceito de afeto com o conceito de *quale*, de modo que processos afetivos são intrinsecamente subjetivos e irreduzíveis a processos cognitivos (para o autor, portanto, o *quale* é irreduzível aos processos cognitivos).

Entretanto, algumas considerações devem ser feitas aqui. A partir de 2010, as publicações de Panksepp ganharam um aspecto mais consensual com a neurociência cognitiva, dado, certamente, as devidas restrições, como veremos adiante. Particular atenção deve ser dada ao seu artigo de 2016, publicado em coautoria com outros três autores, sendo dois cognitivistas e dois afetivistas (Panksepp et al., 2016). Esse artigo possui um formato interessante, cada par de autores defende uma das posições e critica a posição oposta, e, ao final, todos os autores tentam encontrar um meio termo entre as posições teóricas defendidas. Nesse artigo, fica bem claro que Panksepp já não adota mais a visão de que processos afetivos e cognitivos são independentes, mas, mesmo assim, o autor deixa claro que os processos afetivos brutos (*raw affects*) são processos subcorticais que continuam existindo mesmo após procedimentos de decorticação (remoção do córtex) (Panksepp et al., 2016, pág. 188), já que, segundo os Panksepp e Solmons, animais decorticados continuam apresentando alguns comportamentos emocionais básicos (emoções básicas de Panksepp descritas acima). Logo, Panksepp argumenta que processos afetivos são incondicionados (*inatos*), mesmo que

representações cognitivas (processos condicionados pela aprendizagem) possam acompanhar tais processos emocionais (Ibidem, pág. 190). Fazendo uma análise filosófica do modelo de Panksepp, podemos dizer que, para o autor, embora possamos estabelecer relações epistemológicas entre processos afetivos e cognitivos (ou seja, mesmo que possamos utilizar determinados processos cognitivos para explicar determinados processos afetivos), os processos afetivos são ontologicamente distintos dos processos cognitivos.

De modo geral, mesmo que a análise de Panksepp aqui apresentada seja breve, esperamos que ela sirva como uma boa introdução para um assunto muito mais denso do aqui apresentado. O importante é ter em mente que o paradigma afetivista embasa um modelo de consciência distinto do modelo cognitivista de consciência apresentado acima.

A consciência afetiva

O paradigma afetivista tem duas ideias centrais que são muito importantes para a compreensão das ideias de Damásio (que serão analisados no próximo capítulo). O modelo afetivista vincula consciência afetiva com os mecanismos de sobrevivência do cérebro, vinculando, desse modo, a atividade consciente com mecanismos subcorticais, por estes serem antigos no processo evolutivo. Portanto, segundo esse modelo, é possível ter um nível basal de consciência mesmo diante da ausência total de áreas corticais, já que, segundo Panksepp (2005), o afeto (que, segundo o autor, compreende o aspecto qualitativo da experiência) está embutido nas áreas subcorticais que instanciam os processos afetivos. Desse modo, esse modelo defende uma visão *subcortical* de consciência, algo que vai diretamente contra o modelo cognitivista analisado na seção anterior. Seguindo essa mesma linha, já que esse modelo está

embasado numa visão evolucionista de emoção, o paradigma afetivista assumi que existem emoções que estão presentes em todos os animais com determinado nível de complexidade comportamental, chamadas de emoções básicas (Panksepp, 2000, as classifica como espécies naturais). Logo, o que devemos ter em mente é que, segundo o paradigma afetivista: 1) as emoções são processos diretamente vinculados com os processos que mantêm a sobrevivência do organismo (circuitos de sobrevivência); 2) há emoções básicas inatamente inscritas nos mecanismos afetivos subcorticais, chamadas de emoções básicas; 3) os processos cognitivos e afetivos são ontologicamente distintos; e 4) a consciência (afetiva) pode ser instanciada por mecanismos subcorticais mesmo diante da total remoção do córtex cerebral (decorticação).

Tendo em vista os dois paradigmas discutidos nesse capítulo, agora poderemos partir para a análise das ideias centrais de António Damásio (1994, 1999, 2003, 2010).

Capítulo 2

A teoria da consciência de Damásio:

conceitos-chave e fundamentos teóricos

Nosso objetivo aqui é realizar uma análise crítica da obra de Damásio, e não apenas apresentarmos passivamente sua teoria, já que o debate sobre a consciência se dá num contexto de certo embate entre posições, e a análise crítica, metodológica e empírica dessas teorias é a única forma de avançarmos nosso conhecimento sobre esse fenômeno.

As contribuições de Damásio para o debate sobre a consciência possuem tanto um peso filosófico quanto científico, já que, como dito na introdução, Damásio não faz uma mera descrição de dados empíricos. Por exemplo, como procuraremos mostrar adiante, a consciência, para o autor, é, acima de tudo, caracterizada pelo “sentimento” (experiência qualitativa no ato de conhecer). Desse modo, Damásio não trata a consciência apenas como um mero conjunto de imagens cognitivas.

Em linhas gerais, o nosso objetivo aqui é analisar as ideias de Damásio, tanto conceitualmente (como os conceitos se alteraram dentro da teoria do autor, e como eles se relacionam entre si, formando ou não um modelo conceitual) quanto empiricamente (quais as bases empíricas das ideias de Damásio). Além disso, buscaremos demonstrar como que a teoria da emoção, sentimento, *self* e consciência de Damásio estão teoricamente embasadas, em certa medida, tanto no paradigma afetivista quanto no paradigma cognitivista discutidos no capítulo anterior.

Corpo e emoção

Sem dúvida, uma das mais importantes contribuições de Damásio (1994), e provavelmente pelas quais ele é mais lembrado e citado, é a relação entre emoção e neurobiologia, temática que enfrentou séria resistência durante muito tempo no contexto científico, já que o estudo das emoções era tido como um objeto da filosofia, provavelmente, devido a heranças cartesianas e positivistas, como também argumentado por LeDoux (1996) e Panksepp (1998). Segundo Damásio (1994), Descartes considerava que havia uma distinção entre razão e emoção, o que já não é mais aceito diante dos dados neurobiológicos atuais, uma vez que tomada de decisão, por exemplo, é um processo cognitivo no qual as consequências emocionais têm um papel fundamental, o que Damásio discute no seu modelo de marcadores somáticos (Damásio, 1994; Damásio et al., 1996). Damásio defende que as dicotomias mente-corpo e razão-emoção são um resquício da herança cartesiana, e que para entendermos a mente, principalmente as emoções, temos de nos utilizar de dados corporais (Damásio, 1994, pág. xiii e xiv).

De modo geral, os grandes pilares da teoria de Damásio são: as emoções e sentimentos (fenômenos que o autor diferencia um do outro), o *self* (em diferentes níveis) e o corpo (principalmente dados interoceptivos), algo que se mantém em todos os livros do autor. cremos que o modo mais eficiente de abordarmos os princípios básicos da teoria de Damásio é dividi-la nos seguintes tópicos: *emoção e sentimentos* [conceitos que mais se alteram na obra do autor]; o *self*; e *a consciência* (o corpo está presente em todos esses temas), analisando como tais ideias estão inseridas tanto nos paradigmas teóricos discutidos no primeiro capítulo quanto no modelo de emoção de William James apresentado abaixo.

A teoria das emoções e sentimentos

A teoria de Damásio sobre as emoções e sentimentos é fundamentada no modelo de emoção de William James (1884), logo, antes de seguirmos, é importante discutirmos alguns pontos centrais da visão de James sobre a emoção.

James

A análise do influente artigo de William James, *What is an emotion?* (1884)³⁹, nos fornece inúmeras ferramentas teóricas para a nossa discussão. Mas antes de analisarmos esse artigo, é importante ter em mente o paradigma geral das ideias de James (1907) baseado no modelo pragmático de mente. Mas, para o nosso objetivo aqui, não é necessário fazer uma análise aprofundada desse modelo de mente, apenas a ideia geral já será o suficiente. Segundo James, a significância de uma ideia é medida a partir da sua capacidade de proporcionar uma nova experiência e uma nova observação. O autor, então, estabelece uma relação direta entre ideia e *ação*. Segundo James (1907), o conhecimento não se dá através da reflexão passiva de experiências passadas, mas sim através da projeção ativa da experiência no futuro mediada pela ação ativa. É importante ter em vista essa ideia geral para a nossa análise a seguir.

No início do artigo de 1884, James especula quais seriam as áreas cerebrais responsáveis pelos processos emocionais, propondo que ou as emoções são instanciadas por um centro neural especializado [como também defendido por Panksepp, já analisado no primeiro capítulo] ou elas são frutos das mesmas áreas cerebrais responsáveis pelos processos sensoriais e motores (James, 1884, pág. 188). E, já de pronto, James argumenta que a segunda hipótese, a partir da teoria que será apresentada pelo autor, é a mais provável⁴⁰. Segundo James, as “emoções tradicionais”⁴¹, como surpresa,

³⁹ Citado 4905 vezes até o momento.

⁴⁰ Pois, como foi dito logo acima, a aplicação de uma ideia, segundo James, só ocorre através da ação.

⁴¹ Podemos dizer que esse é o termo utilizado pelo autor para se referir ao que hoje se chama “emoções básicas”.

curiosidade, êxtase, medo, raiva, luxúria, ambição e gosto, são caracterizadas por expressões corporais que são, de certo modo, compartilhadas pela maioria das pessoas. A ideia mais famosa desse artigo é que, segundo James, as reações corporais associadas com essas emoções tradicionais não se originam a partir de um estado mental afetivo, mas sim o oposto, são as reações corporais que dão origem aos estados mentais afetivos (Ibidem, pág. 189 e 190). Como o próprio autor coloca: “(...) só sentimos culpa porque nós choramos, raiva porque nós atacamos, medo porque nós trememos (...)” (Ibidem, pág. 190, tradução nossa).

Como justificativa para essa tese, James recorre à visão evolucionista sobre o funcionamento do sistema nervoso, vinculando as reações emocionais com os mecanismos de sobrevivência. Segundo o autor, o sistema nervoso é caracterizado por um conjunto de predisposições. Para James, o modo como nos comportamos emocionalmente, como quando amamos, sentimos raiva, etc., ocorre de maneira muito semelhante entre os indivíduos da nossa espécie, já que, segundo o autor, os trabalhos de Darwin e sucessores demonstram que essas formas de se comportar estão estampadas no sistema nervoso de cada espécie (Ibidem, pág. 190 e 191). Logo, assim como Panksepp, analisado no capítulo anterior, James estabelece uma ligação entre reações emocionais e circuitos de sobrevivência herdados no processo evolutivo.

James chega a declarar que sem as reações emocionais, só sobraria um estado intelectual frio e neutro (Ibidem, pág. 193). Como o autor mesmo coloca: “uma emoção humana desincorporada pura é uma não-entidade” (Idem, pág. 194, tradução nossa). De modo geral, é interessante notar que, no decorrer desse artigo, James deixa claro que ele dissocia os processos emocionais dos processos cognitivos⁴²: “sem os estados corporais que seguem a percepção, o último [a percepção] seria *puramente cognitivo* em forma,

⁴² Assim como os afetivistas clássicos, como Zajonc (1984), posteriormente defenderam.

pálido, incolor, destituído de calor emocional” (Ibidem, pág. 190, tradução nossa, itálico nosso); “(...) se eu tivesse de me tornar corporalmente anestésico, eu seria excluído da vida das afeições, severas e ternas, e arrastaria uma existência de uma forma *meramente cognitiva ou intelectual*” (Ibidem, pág. 194, tradução nossa, itálico nosso). Logo, vê-se nessa obra de James tanto as dicotomias emoção-intelecto e cognição-corpo quanto a dicotomia emoção-cognição, e todas essas dicotomias não são toleradas no contexto neurobiológico atual (Damásio, 1994).

De qualquer forma, tanto a relação entre emoções e corpo quanto a relação direta entre emoção e sobrevivência propostas por James (1884) são fundamentais para entendermos o modelo de emoção e sentimento de António Damásio.

Emoção e sentimento segundo Damásio

Já no livro de 1994, *Descartes' Error*, Damásio declara que sua teoria da emoção é baseada na teoria de James (1884), porém, Damásio (1994, pág. 129; 2010, pág. 80) afirma que, por mais brilhante que a teoria de James seja, este autor considerava que a reação emocional seguia simplesmente o modelo de estímulo-resposta, algo que serve, segundo Damásio (1994), para o que é chamado de emoções primárias, mas não é o suficiente para explicar emoções mais complexas, como as emoções sociais, por estas possuírem um fator sociocultural mais desenvolvido. Para Damásio (1994, pág. 131), emoções primárias são reações fisiológicas que ocorrem contingentemente à imediata apresentação do estímulo, sem que o sujeito tenha de raciocinar ou refletir sobre a situação na qual o estímulo está inserido. O maior e mais investigado exemplo desse tipo de reação é o “medo”, cujas reações comportamentais, incluindo a expressão facial, ocorrem imediatamente após a apresentação do estímulo. Tanto que alguns estímulos que causam reações de medo são inatos, como aranhas e

cobras [ver o exemplo do galho retorcido de LeDoux (2015) no próximo capítulo]. Podemos dizer que esse tipo de emoção ocorre segundo o modelo $S \rightarrow R$ (estímulo-resposta) do condicionamento respondente (modelo base na investigação das reações de medo até hoje). Segundo Damásio (1994, pág. 133)⁴³, as bases neurais das emoções primárias são a amígdala [importantíssima estrutura para as reações de medo, mas não a única] e o córtex cingulado anterior [segundo o famoso artigo de Damásio et al. (2000), essa área está correlacionada com felicidade e raiva].

Diferentemente das emoções primárias, as emoções secundárias dependem de avaliação cognitiva e consciente sobre a situação na qual o estímulo está inserido, por exemplo, as reações emocionais em relação à morte de uma pessoa desconhecida são completamente distintas das de um ente querido, e isso demanda avaliação cognitiva da informação. Para o autor, esse tipo de emoção depende da participação do córtex pré-frontal, principalmente sua área ventromedial (Bechara et al., 1994)⁴⁴, já que esse tipo de reação, segundo Bechara et al. (1994), se baseia em representações corporais adquiridas durante o desenvolvimento ao invés de serem inatas [o uso de reações emocionais prévias em processos cognitivos, como tomada de decisão, é um ponto importante dentro da teoria do autor], necessitando inclusive de processos de tomada de decisão, uma vez que, como já dito, tais emoções dependem de avaliação cognitiva da situação. E, segundo Damásio (1994, pág. 137), após o processamento pré-frontal da situação, essa área envia informações à amígdala e ao córtex cingulado anterior, dando início à reação emocional. No livro de 1994, Damásio trata as reações emocionais como inconscientes, algo que será aprofundado no livro de 1999, *The feeling of what happens*.

⁴³ Lembrando que nessa época as pesquisas sobre reações emocionais ainda estavam no início, já que é um objeto de estudo recente na neurobiologia.

⁴⁴ Nesse artigo, Bechara, Damásio e colegas debatem dados neurológicos sobre pacientes que perderam sensibilidade às consequências de longo prazo devido lesões no córtex pré-frontal ventromedial. O papel dessa área na tomada de decisão é também discutido por Bechara, Damásio e colegas em um artigo posterior (Bechara et al., 1999)), e Bechara, Tranel e Hanna Damásio publicaram outro artigo sobre o tema em 2000. Esses três artigos são bem reconhecidos na literatura.

Sobre os sentimentos, Damásio (1994, pág. 143) defende que esses fenômenos são distintos das emoções, mesmo que vinculados. O autor propõe que a maioria dos sentimentos parte de reações emocionais, o que ele chama de *sentimento emocional*. Mas, segundo o autor, há sentimentos que não se originam desse tipo de reações, o que ele chama de *sentimento de fundo*. Os sentimentos emocionais, seguindo mais ou menos as ideias de James (1884), são a percepção do que ocorre no corpo durante um contexto emocional, através de informações intero e proprioceptivas (como o músculo da face) enviadas aos córtices somatossensoriais, parietais e insulares.

Os sentimentos de fundo, por sua vez, segundo Damásio (Ibidem, pág. 150), surgem a partir das representações do corpo processadas pelos córtices somatossensoriais, parietais e insulares, sem que seja necessário ocorrer uma reação emocional. Podemos dizer que o sentimento de fundo é ligado à corporeidade em si, ou seja, ele emerge diretamente do corpo vivo, sem a necessidade de ser mediado pela representação de um estímulo externo, como no caso das emoções e sentimentos emocionais. Para Damásio (1994, pág. 153), pacientes com anosognosia perdem a capacidade de produzir representações do estado corrente do corpo, logo, eles não conseguem perceber suas próprias limitações decorrentes de determinada doença, ou seja, eles não conseguem atualizar os mapas do corpo, e, assim, o sentimento de fundo, sendo a base para as reações sentimentais, fica também desatualizado⁴⁵.

De modo geral, um dos pontos fundamentais da teoria é o papel do corpo para as reações emocionais, pois tanto as emoções quanto os sentimentos se originam no corpo, e, como o autor mesmo afirma, têm o corpo como teatro (Ibidem, pág. 152).

Mas, definitivamente, uma das ideias mais brilhantes de Damásio (1994) foi o que o autor chama de *marcador somático*, hipótese que Damásio e colaboradores

⁴⁵ Já nessa época podemos observar as raízes da teoria de Damásio sobre a consciência, já que, para o autor, a percepção daquilo que ocorre é em si um sentimento.

aprofundam anos depois (Damásio et al., 1996). A partir dessa hipótese, Damásio postula que estados corporais, presentes ou prévios, influenciam na tomada de decisão, ou seja, experiências emocionais prévias ou atuais [que ocorrem necessariamente no corpo] influenciam na avaliação cognitiva de possíveis consequências para uma determinada ação. E é devido a isso que o indivíduo avalia determinada ação como ruim ou boa: pelas consequências emocionais/orgânicas que ela possivelmente causará. Desse modo, o autor quebra a dicotomia entre cognição e emoção que vinha desde as teorias de Descartes.

Em suma, os marcadores somáticos são uma instância especial de sentimentos gerados a partir das emoções secundárias. Essas emoções e sentimentos foram conectados, através do aprendizado, aos resultados futuros previstos a partir de certos cenários (Damásio, 1994, pág. 174, tradução nossa, itálico do autor).

E é interessante que já naquela época Damásio e colaboradores (1996, pág. 1414) colocavam o córtex pré-frontal como uma importante área para essa função com base em um conjunto de dados sobre lesões no córtex ventromedial. E, hoje em dia, o controle top-down do córtex pré-frontal sobre as reações emocionais [como o controle do córtex pré-frontal lateral sobre a amígdala central, por exemplo] já está estabelecido na literatura (Etkin e Wager, 2007).

Tomando agora o livro de 1999 (pág. 50 e 51), *The feeling of what happens*, Damásio divide o conceito de emoção em três tipos: a) as emoções básicas (alegria, tristeza, medo, raiva, surpresa e nojo), tidas como inatas e universais, que ocorrem imediatamente à apresentação do estímulo e são reconhecidas por diferentes culturas; b) as emoções aprendidas e inseridas num contexto cultural, como inveja, culpa e orgulho; c) e as emoções de fundo, caracterizadas por modularem as demais emoções, como uma espécie de humor, como calma, tensão, fadiga e bem-estar, decorrentes do estado corporal corrente.

Nessa obra, o autor consolida a ideia de que as emoções são primariamente criadas por áreas subcorticais (Ibidem, pág. 61 e 62), como o tronco cerebral⁴⁶, o hipotálamo e a amígdala, além de áreas corticais, como o já mencionado córtex pré-frontal ventromedial. Como pode ser observado, o autor retira a ínsula e o córtex somatossensorial como áreas responsáveis por reações afetivas [o que é retomado em seu livro de 2003], colocando essas áreas apenas como responsáveis pelos mapas do corpo [base de sua teoria sobre o *self*, como veremos posteriormente]. E é importante ressaltar que a partir daqui, Damásio classifica claramente emoção como um processo plenamente inconsciente (reformulando, assim, a teoria de 1994) [algo que se mantém em todos os trabalhos do autor]. Portanto, todas as reações bioquímicas e suas consequências comportamentais, como postura corporal e expressões faciais, são processadas e moduladas sem a percepção consciente do indivíduo, principalmente se for uma das emoções básicas descritas acima. O controle de um processo emocional só é possível quando este passa a ser um sentimento, ou seja, quando as reações corporais decorrentes de um estado emocional são experienciadas conscientemente pelo indivíduo ou quando a própria reação emocional é inibida inconscientemente pelo córtex pré-frontal (Ibidem, pág. 57).

Sobre os sentimentos, é interessante, e também controverso, o fato de o autor os dividir em dois tipos: sentimentos inconscientes e sentimentos conscientes (Damásio, 1999, pág. 170). Sentimentos conscientes são experiências conscientes decorrentes das mudanças corporais causadas por reações emocionais; e sentimentos inconscientes são reações emocionais que fogem à percepção direta do indivíduo, por exemplo, alguém que está tenso, inclusive tremendo uma das pernas, mas não percebe conscientemente

⁴⁶ Posição extremamente controversa hoje em dia, algo que, além do que discutido acima, vamos aprofundar no final dessa seção.

tal estado corporal. Podemos dizer que esse tipo de sentimento é uma *experiência inconsciente*⁴⁷ de uma emoção de fundo.

Por sua vez, no livro posterior de Damásio, *Looking for Spinoza* (2003), mesmo o autor não alterando sua teoria sobre emoções e sentimentos do livro de 1999 [mantendo a divisão de sentimento entre inconsciente e consciente], Damásio (2003) atualiza alguns pontos importantes sobre a base neurobiológica desses fenômenos [que é a parte mais interessante desse livro para a presente dissertação], onde o autor emprega os dados coletados e publicados em seu famoso artigo de 2000 (Damásio et al., 2000), no qual o Damásio e colaboradores utilizaram tomografia por emissão de pósitrons (PET) para investigar os correlatos neurais de emoções autogeradas em 41 participantes. Os resultados dessa pesquisa são citados até hoje pela literatura médica e neurocientífica⁴⁸ por ser um dos primeiros e mais acurados mapas neurais de quatro das seis emoções básicas mencionadas acima, no caso, tristeza, alegria, raiva e medo.

⁴⁷ Como o autor a classifica, mas nós reconhecemos que é uma definição problemática.

⁴⁸ Citado 2098 vezes até a escrita dessa dissertação.

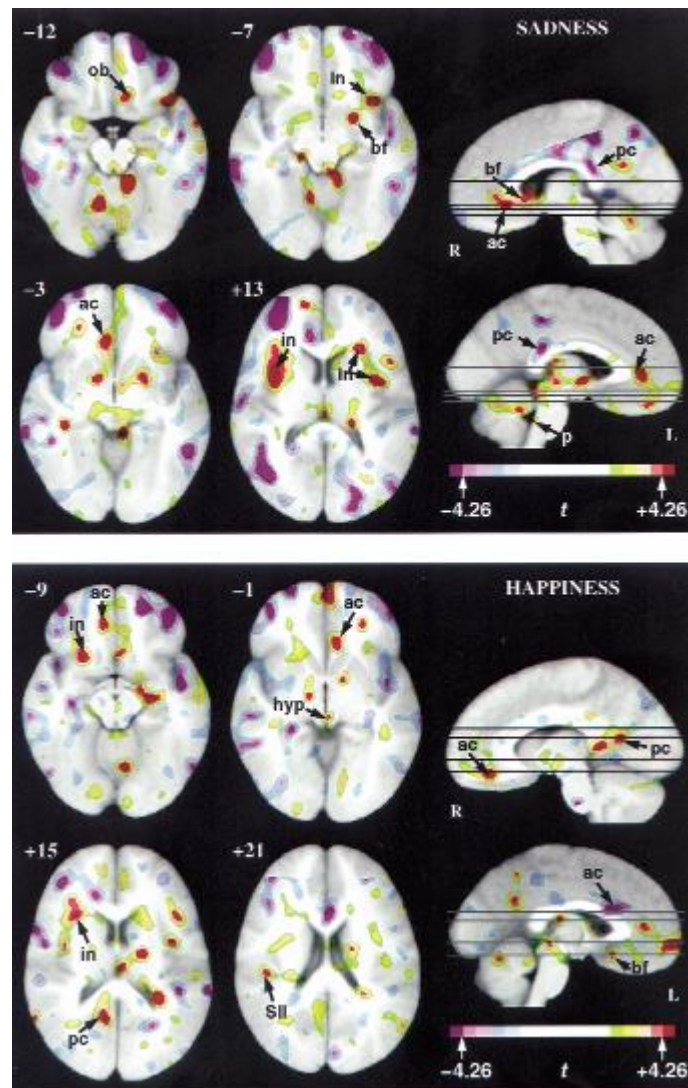


Fig. 11. Áreas neurais correlacionadas com tristeza e felicidade. Zonas avermelhadas denotam maior ativação enquanto zonas roxas denotam maior desativação. (Ilustração retirada de Damásio et al., 2000, pág. 1050).

Na figura acima, podemos observar uma grande distinção entre os picos de ativação e desativação entre as duas emoções. Tristeza está correlacionada com ativação bilateral da ínsula e do córtex cingulado anterior, além de ativação na ponte (*pons*) dorsal, ao mesmo tempo em que ocorre significava desativação do córtex somatossensorial posterior. Já quando os participantes relatam alegria, nota-se maior ativação do córtex cingulado posterior, na ínsula esquerda e no córtex somatossensorial esquerdo, além de desativação significativa no córtex cingulado anterior esquerdo.

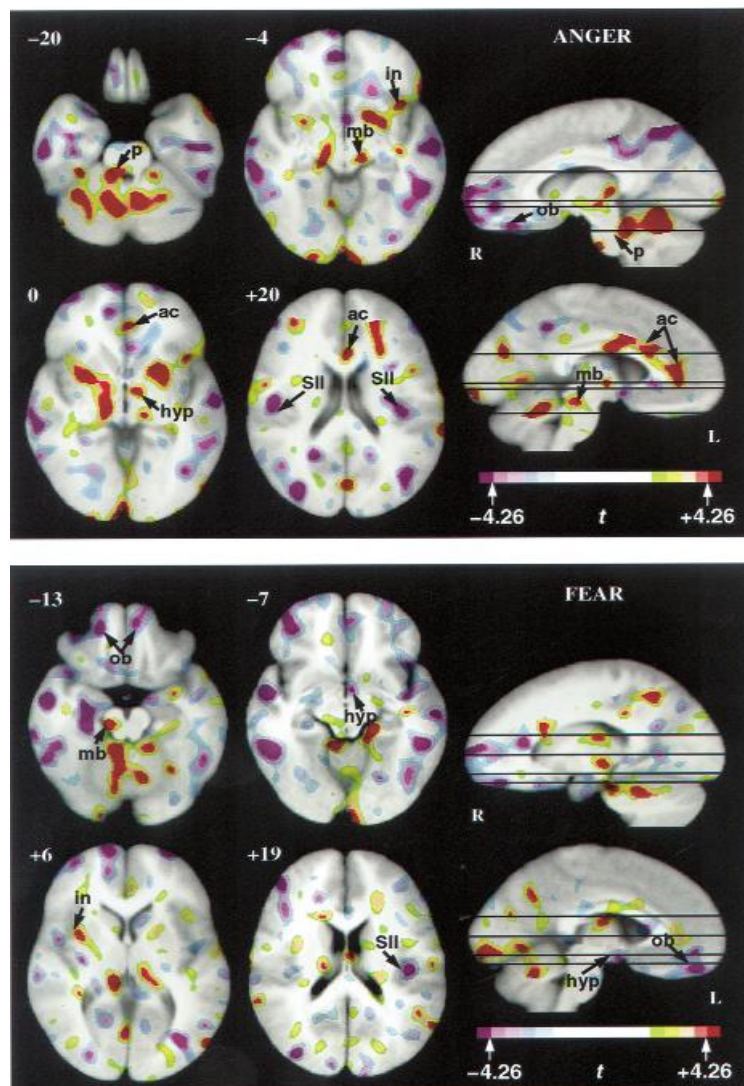


Fig. 12. Áreas neurais correlacionadas com tristeza e felicidade. (Ilustração retirada de Damásio et al., 2000, pág. 1051).

Já na figura acima, podemos observar que raiva está correlacionada com ativações no mesencéfalo, na ponte e na metade do córtex cingulado anterior esquerdo, além de desativação em ambas as partes do córtex somatossensorial e no córtex orbitofrontal; já em relação ao medo, podemos observar ativação no mesencéfalo e desativação no córtex somatossensorial esquerdo, hipotálamo e córtex orbitofrontal.

A partir desses trabalhos, a importância da ínsula e do tronco cerebral superior passa a ser um ponto-chave na teoria dos sentimentos e *self* de Damásio, principalmente em seu livro mais recente (Damásio, 2010).

Por sua vez, no livro *The Self Comes to Mind*, Damásio (2010) conceitua as emoções como “programas de ação desenvolvidos pela evolução” (Ibidem, pág. 83, tradução nossa), caracterizando melhor os aspectos comportamentais desses fenômenos, já que emoções são reações bioquímicas que alteram as vísceras, a propriocepção, a postura do corpo e a expressão facial, além de modularem os aspectos cognitivos momentaneamente, fazendo com que alguns pensamentos e ações sejam muito mais prováveis que outros, por exemplo, quando diante de um predador é muito mais provável pensarmos em como e para onde fugiremos, ou como nos defenderemos, ao invés de pensarmos em sexo ou sobre a viagem do mês que vem.

Sentimento emocional, por outro lado, é a *percepção* daquilo que ocorre no nosso corpo e nossa mente durante um estado emocional (Damásio, 2010, pág. 76). Para Damásio (2010), sentimentos são imagens de ações ao invés de ações propriamente ditas. Como afirmado pelo próprio autor, sentimentos são percepções realizadas nos próprios mapas cerebrais (Idem, ibidem). E é importante ressaltar que os mapas cerebrais do corpo⁴⁹, mesmo possuindo uma base estável⁵⁰, se alteram com o tempo dependendo, principalmente, do estado visceral do organismo, desse modo, a *interocepção* toma um lugar muito importante na teoria de Damásio sobre os sentimentos.

Uma das grandes influências de Damásio são claramente os famosos artigos de Arthur D. Craig (2002, 2003, e, principalmente, 2009). Nesses trabalhos, Craig desenvolve uma importante e influente teoria sobre a interocepção e seu papel nas reações afetivas e consciência humana, atribuindo um papel crucial à ínsula anterior. Craig (2009) defende que o córtex insular anterior está relacionado com controle voluntário, ilusão da mão de borracha (no qual o cérebro interpreta um membro externo

⁴⁹ Trataremos isso melhor na teoria sobre o *self*.

⁵⁰ O *protossself*.

como pertencente ao próprio corpo), auto-reconhecimento, percepção do tempo e atenção. Isso, obviamente, acrescentado ao papel da ínsula nas reações emocionais discutidas no artigo de Damásio et al. (2000), brevemente apresentado acima.

Para Damásio (2010, pág. 88), a reação emocional é uma espécie de descarga bioquímica que modifica toda a paisagem do corpo. Durante uma reação de medo, por exemplo, ocorre grande ativação da amígdala e do hipotálamo, alterando a frequência dos batimentos cardíacos (dependendo da reação de medo tomada) e a pressão sanguínea, cortisol é secretado no sangue e a expressão facial se altera se adequando às características de medo. Um ponto importante é o papel que Damásio atribui ao tronco cerebral, principalmente à substância cinzenta periaquedutal (PAG), na modulação das respostas corporais. Desde o artigo de Bandler e Shipley (1994), a PAG vem sendo relacionada com respostas defensivas, analgesia, regulação autonômica e reações emocionais negativas, como estresse, dor e medo. Desse modo, Damásio (2010) argumenta que a PAG é responsável por selecionar qual resposta de medo será empregada, no caso, correr ou congelar (*freezing*), enviando, posteriormente, informações ao cerebelo, responsável por iniciar a ação.

Já os sentimentos emocionais ocorrem quando essas alterações são percebidas através das informações corporais enviadas aos córtices somatossensoriais e à ínsula (Ibidem, pág. 83 e 84). Segundo o autor, essas áreas estão relacionadas com a produção dos mapas neurais do corpo, e, como dito acima, os sentimentos emocionais são criados nesses mapas. Desse modo, quando ocorrem alterações na paisagem do corpo devido a reações emocionais, ocorrem também alterações nos mapas do estado corrente do corpo (informação próprio e interoceptiva), dando então origem à percepção do que está ocorrendo no corpo durante um estado emocional.

Essas representações estáveis do corpo são o que Damásio chama de sentimentos primordiais. Para ele, esses mapas geram um sentimento base para todos os sentimentos emocionais: um sentimento ligado à corporeidade em si (Damásio, 2010, pág. 70). E, para o autor, esses sentimentos são criados pelos núcleos do tronco cerebral superior e pela massa cinzenta periaquedutal (PAG), pois são áreas que recebem informações corporais diretamente, podendo tomar ações sem a necessidade do processamento cortical (Ibidem, pág. 54). O ponto é o seguinte: mesmo que essas áreas possuam muitos mecanismos de ação automatizados (coisa que Damásio só concorda em parte), se elas são capazes de iniciar a ação é porque, dentro da teoria de Damásio, elas produzem mapas do corpo. No caso, já que a PAG pode iniciar o comportamento de fuga isso se deve ao fato de que essa região cria mapas do corpo pelos quais tal ação é desencadeada.

É importante ressaltar que nessa obra Damásio deixa de considerar sentimentos emocionais como inconscientes, como na obra de 1999, onde o autor dividia os sentimentos emocionais em conscientes e inconscientes. Entretanto, o autor não deixa claro se os sentimentos primordiais são inconscientes ou pré-conscientes. De acordo com o autor, os sentimentos primordiais são “(...) reflexões espontâneas do estado do corpo vivo” que (...) “resultam de nada além do corpo vivo e precedem qualquer interação entre a regulação da maquinaria da vida e qualquer objeto” (Damásio, 2010, pág. 101). Então, como eles resultam da (sensação) do corpo vivo e são a base de todos os sentimentos emocionais, é possível que eles estejam em um nível subconsciente da mente.

Tendo em mente as informações apresentadas acima, possuímos, agora, ferramentas argumentativas para analisarmos os fundamentos teóricos da teoria das emoções e sentimentos de Damásio.

Fundamentos teóricos da teoria da emoção e sentimento de Damásio

A teoria de Damásio sobre as emoções e sentimentos é um ótimo exemplo para observarmos como os pontos centrais de cada paradigma da consciência analisados no primeiro capítulo, juntamente com o modelo de James analisado acima, influenciam na teoria do autor.

Assim como James e Panksepp, Damásio mantém o vínculo entre emoção e mecanismos de sobrevivência, atribuindo emoção à atividade de mecanismos subcorticais que ocorre previamente ao processamento cognitivo amplo e consciente do estímulo. Como o autor coloca: “as emoções são curiosas adaptações que são parte e parcela da maquinaria com a qual os organismos regulam a sobrevivência. Antigas como as emoções são na evolução, elas são um componente de alto nível dos mecanismos de regulação da vida” (Damasio, 1999, pág. 51, tradução nossa). Damásio (2010), como visto acima, chega a classificar emoção como um programa de *ação*, deixando claro o modelo pragmático de James em sua teoria. Outro fator importante é notar que, para Damásio, o sentimento é caracterizado pela experiência qualitativa de uma emoção (ou, como dito acima, experiência qualitativa no ato de conhecer). Para o autor, a emoção é caracterizada por reações observáveis (tanto através das expressões corporais e faciais quanto através das mudanças fisiológicas associadas a processos emocionais), enquanto o sentimento é uma experiência mental privada e não-observável da emoção:

Eu tenho proposto que o termo sentimento deve ser reservado para a experiência mental privada de uma emoção, enquanto o termo emoção deve ser usado para designar a coleção de respostas, muitas das quais são publicamente observáveis. Em termos práticos, isso significa que você não pode observar um sentimento em outra pessoa, embora você possa observar um sentimento em si mesmo quando, como um ser consciente, você percebe seus próprios estados emocionais (Damásio, 1999, pág. 39, tradução nossa).

Assim como Panksepp (1998, 2005), Damásio considera que a experiência afetiva é algo subjetivo, porém, Panksepp atribui essa experiência subjetiva ao que ele chama de afeto, e, ao contrário de Damásio, Panksepp não defende um modelo inconsciente de afeto/emoção. Como visto no primeiro capítulo, Panksepp classifica afeto como um sentimento experiencial subjetivo que não tem como ser transmitido através da linguagem (Panksepp, 2005, pág. 32). E o conceito que Panksepp (2005) dá às emoções é realmente próximo do conceito de Damásio, já que o autor considera que emoções são “afetos direcionados para a ação” (Idem, *ibidem*). Porém, emoção, para Panksepp, é um fenômeno afetivo, e, já que o autor atribui os aspectos subjetivos da experiência ao afeto, toda emoção, mesmo sendo direcionada à ação, é baseada em um aspecto afetivo subjetivo. Logo, a distinção entre afeto, emoção e sentimento não faria sentido a partir do modelo de Panksepp, já que esse autor atribui as características qualitativas ao afeto, diferentemente de Damásio, que as atribui ao sentimento (classificando emoção como uma reação inconsciente). De qualquer forma, atribuir aspectos qualitativos ao afeto é fundamental na teoria de Panksepp, já que, como analisado no primeiro capítulo, o autor defende que estruturas subcorticais podem instanciar reações afetivas sem a necessidade de processos corticais.

Mas, o mais importante aqui é analisar as diferenças sobre a inclusão (ou exclusão) de processos cognitivos nos modelos de emoção de Damásio, Panksepp e James. Como visto acima, James deixa claro uma distinção entre processos corporais emocionais e processos intelectuais cognitivos, pelo qual não haveria processos afetivos/emocionais apenas através de processos cognitivos. E, mesmo que a teoria de James tenha influenciado fortemente o modelo de Damásio, este último possui uma visão bem diferente sobre a influência de processos cognitivos nos processos emocionais e sentimentais. Segundo Damásio (2010), as emoções são programas de

ação, mas essas ações “(...) são complementadas por um programa cognitivo que inclui certas ideias e modos de cognição (...)” (Damásio, 2010, pág. 76), logo, mesmo que o autor considere que as emoções são programas de ação largamente automáticos, Damásio não desvincula nem ignora a importância de processos cognitivos para as reações emocionais. E isso fica ainda mais claro quando o autor discute a natureza dos sentimentos, já que, como analisado acima, os sentimentos são caracterizados pela percepção do corpo físico através dos mapas do corpo durante estados emocionais. Segundo Damásio (2010), “(...) os sentimentos são *imagens de ações* ao invés de ações propriamente ditas; o mundo dos sentimentos é constituído pelas percepções executadas nos mapas do cérebro” (Ibidem, pág. 76, tradução nossa, itálico nosso). Logo, Damásio (2010) deixa visível a importância das imagens cognitivas na construção de processos sentimentais. Já Panksepp possui uma posição um tanto quanto distinta da de Damásio.

Panksepp (2004) declara que:

(...) enquanto a psicologia e a neurociência permanecerem mais preocupadas com os impressionantes sistemas cortico-cognitivos do cérebro humano do que com os afetivos subcorticais, nossa compreensão das fontes da consciência humana permanecerá lamentavelmente incompleta. (Panksepp, 2004, pág., 58).

Logo, Panksepp, pelo menos até 2009, deixava bem claro a distinção entre processos corticais-cognitivos e subcorticais-afetivos. E, mesmo que o autor tenha recentemente feito uma reaproximação com a visão cognitivista das emoções (Panksepp et al., 2016), Panksepp, como vimos no primeiro capítulo, continua considerando uma distinção ontológica entre processos afetivos e cognitivos, de modo que o autor argumenta que os processos cognitivos são caracterizados por representações condicionadas (aprendidas) em contrapartida dos processos afetivos, os quais são caracterizados por representações afetivas cruas *inatas* (representações incondicionadas).

Portanto, embora Damásio compartilhe algumas ideias centrais com James e Panksepp, sua teoria não dialoga apenas com as teorias desses últimos. Damásio, na verdade, continua embasando sua teoria em uma visão cognitivista de emoção e sentimento, onde representações cognitivas possuem um papel fundamental na construção dos estados emocionais.

E, mesmo que as reações emocionais sejam fundamentais para compreendermos a teoria da consciência proposta por Damásio, não seria possível, segundo o autor, haver sentimento e, logo, consciência, sem que as reações emocionais sejam direcionadas a um *self*, aspecto que consideraremos a seguir.

A teoria do *self*

Mesmo que a teoria base sobre o *self* no modelo teórico de Damásio tenha sofrido algumas alterações do livro de 1999, *The feeling of what happens*, para o de 2010, *The self comes to mind*, o autor manteve os conceitos e teorias centrais do seu modelo teórico sobre esse fenômeno. No caso, as únicas mudanças teóricas relevantes realizadas são a subdivisão do *protoself* em mapas mestres e a atualização da fundamentação neurobiológica dos níveis de *self*. Desse modo, cremos que seja melhor e mais econômico apenas abordarmos, mesmo que brevemente, sua teoria sobre os três níveis do *self* apenas a partir do livro de 2010 para que possamos, então, analisar os pontos centrais de sua teoria da consciência.

O primeiro nível do *self*, tanto na teoria de 2010 quanto na de 1999, é o *protoself*. Segundo Damásio (2010, pág. 133 e 134), esse nível do *self* é uma coleção integrativa de processos neurais que mapeiam continuamente os aspectos mais estáveis do organismo, no caso, as vísceras. As vísceras, como argumenta o autor (Ibidem, pág. 137), são as áreas mais estáveis do corpo, já que a musculatura e os ossos tendem a

aumentar de tamanho no decorrer do desenvolvimento, além de poderem sofrer outras modificações, por exemplo, quando alguém pratica algum esporte profissionalmente. Mesmo assim, Damásio (2010) defende que há três tipos de mapas que constituem o *protossself*: os mapas mestres interoceptivos, baseados em informação visceral, os quais são produzidos tanto por áreas subcorticais, como os núcleos do trato solitário, núcleos parabraquiais, a PAG, o hipotálamo e o colículo superior, quanto por áreas corticais, como o córtex insular e o córtex cingulado anterior (Ibidem, pág. 134 a 138); os mapas mestres do organismo, que representam o esquema corporal completo, incluindo cabeça, tronco e membros (Damásio não cita nenhuma região neural específica vinculada a esses mapas) (Ibidem, pág. 138); e os mapas dos portais exteriormente direcionados, constituídos por informações advindas das áreas sensoriais, processados principalmente pelos córtices somatossensoriais (Ibidem, pág. 139 a 142). Esse mesmo conjunto de mapas dá origem aos sentimentos primordiais (Ibidem, pág. 142), que é o sentimento de corporeidade, que dão origem a todos os sentimentos emocionais (Ibidem, pág. 70).

Segundo o autor:

Não se pode explicar completamente a subjetividade sem saber sobre a origem dos sentimentos e reconhecer a existência dos sentimentos primordiais, [que são] reflexões espontâneas do estado do corpo vivo. Na minha opinião, os sentimentos primordiais não são mais do que o corpo vivo e precedem qualquer interação entre o mecanismo de regulação da vida e qualquer objeto. Os sentimentos primordiais são baseados na operação dos núcleos do tronco superior do cérebro, que são parte integrante da maquinaria de regulação da vida. Os sentimentos primordiais são a base de todos os outros sentimentos (Damásio, 2010, pág. 70, tradução nossa).

O segundo nível do *self* é o *self central*, e é a partir desse nível, de acordo com a teoria de Damásio (1999, 2010), que é possível ter atividade consciente. A base neural é a mesma mencionada acima (para o *protossself*), na verdade a ideia base do *self central* é bem simples: o organismo interage com o mundo através do corpo (obviamente), desse modo, inúmeros estímulos externos alteram a paisagem estrutural e bioquímica do

corpo, e essa alteração é contraposta aos mapas estáveis do corpo citados acima, o que imediatamente ocasiona o processamento de que *algo está acontecendo* no corpo (Damásio, 2010, pág. 142 e 143). Ou seja, cada pulsão de *self* central é baseada no resultado de alterações nos mapas estáveis do *protoself*, causando o que Damásio (2010) chama de *sentimento de conhecer* (Ibidem, pág. 143 e 144), já que, quando o organismo interage com o objeto (ou estímulo), este é processado como algo que não pertence ao corpo, causando mudanças nos três tipos de mapas do *protoself*, uma vez que a postura corporal se altera, como a posição dos membros e cabeça, além de eliciar atividade nos órgãos sensoriais, o que pode alterar em diferentes níveis a estrutura visceral do organismo (como o estímulo “cobra”, desencadeando uma reação de medo, por exemplo). Através dessa interação, o organismo conhece as peculiaridades do objeto (a partir das reações corporais que ele causa), gerando, em contrapartida, uma diferença entre o organismo e o objeto, dando origem, portanto, a um sentimento de propriedade do próprio corpo.

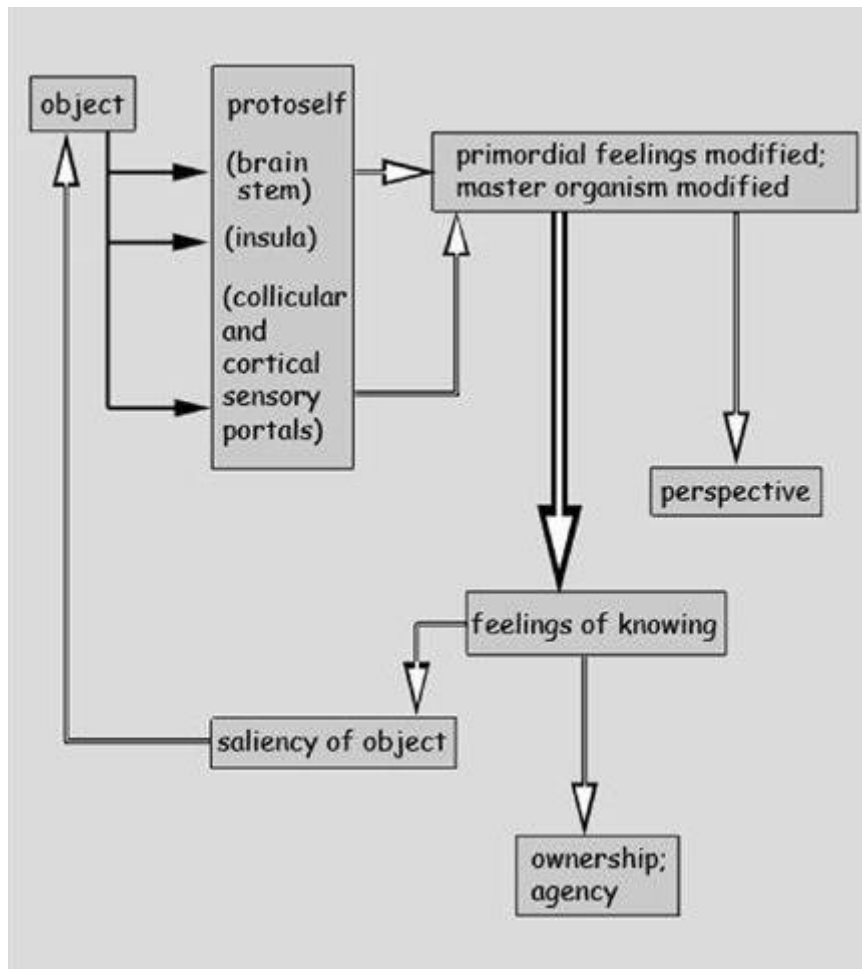


Fig. 13. Mecanismos que dão origem ao *self* central. (Ilustração tirada de Damásio, 2010, pág. 146).

O terceiro nível do *self* é o *self* autobiográfico. Podemos dizer que esse é o nível mais importante no caso da experiência consciente humana, já que esse tipo de *self* é constituído por memórias de primeira pessoa, sendo a base fundamental da autoconsciência (Damásio, 2010, pág. 149). O esquema teórico aqui também é simples (Ibidem, pág. 150 e 151): organismos com um alto nível de complexidade neural possuem mecanismos pelos quais experiências passadas (interações prévias com o ambiente) podem ser usadas como base para ações futuras através de mecanismos mnemônicos. Logo, tanto informações prévias sobre os objetos quanto informações sobre o próprio organismo nessa interação (o que ocorreu no corpo durante a interação) podem ser mantidas através desses mecanismos mnemônicos, ou seja, interações

passadas podem continuar interagindo e modificando o corpo, constituindo um *self* que possui mecanismos para olhar para o passado e se projetar no futuro.

Um fator importante é que Damásio (2010, pág. 153) reconhece que, para tanto, é necessário possuir mecanismos neurais que integrem e coordenem diferentes tipos de processamentos, como alterações corporais e sensoriais prévias e atuais, dando origem a um *self* capaz de alterar sua interação com o ambiente com base em interações prévias e usar os resultados interação atual em interações futuras. Assim, Damásio (2010, pág. 154 e 155) retoma um estudo realizado sobre os córtices posteromediais (Parvizi et al., 2006), o qual atesta que essa região possui um grande número de conexões com inúmeras áreas cerebrais relacionadas com processamento corporal, sensorial e mnemônico, como o córtex cingulado anterior, o córtex pré-frontal medial dorsolateral, os córtex pré-frontais laterais, o claustró, o prosencéfalo basal e as regiões *parahipocampais* (grandemente relacionadas com processamento mnemônico). Desse modo, Damásio (2010) defende que os córtices posteromediais são a possível base neural para o *self* autobiográfico.

(...) os cortices posteromediais (CPMs) parecem ter uma hierarquia funcional maior (...) e exibe vários traços anatômicos e funcionais que os distinguem (...). As evidências obtidas nos últimos anos sugerem que a região dos CPM está realmente envolvida na consciência, especificamente em processos auto-relacionados, e têm fornecido informações anteriormente indisponíveis sobre neuroanatomia e fisiologia da região (Damásio, 2010, pág. 154, tradução nossa).

É interessante notar que, como visto no primeiro capítulo, Tononi et al. (2016) e Koch et al. (2016) também defendem que os córtices posteriores são as principais áreas responsáveis pela consciência, sendo, portanto, um ponto em comum entre a teoria de Damásio (2010) e algumas das hipóteses mais recentes sobre esse fenômeno.

Agora nos resta analisar sobre que a teoria de Damásio sobre o *self* está fundamentada.

Fundamentos da teoria de Damásio sobre self

Como pudemos ver acima, William James (1884) considera que as mudanças corporais ocorrem antes da experiência consciente da emoção. Mas como isso seria possível? Como seria possível o corpo reagir a um estímulo antes de uma avaliação cognitiva ampla e antes da experiência consciente do estímulo e das mudanças corporais? A nosso ver, a teoria do *self* de Damásio se baseia nesse ponto.

Como vimos na seção “fundamentos da teoria da emoção e sentimento” acima, Damásio, ao mesmo tempo em que compartilha a ideia de James e Panksepp que as reações emocionais são processos ligados tanto à sobrevivência quanto ao aspecto qualitativo da mente, o autor não estabelece uma distinção nítida entre processos emocionais e cognitivos (na verdade, como vimos, o modelo de emoção e sentimento de Damásio é baseado em processos cognitivos). Cremos que o mesmo ocorre com a sua teoria do *self*.

Assim como na teoria da emoção, Damásio deixa clara a importância do corpo na construção do *self*. Como pudemos ver, o autor atribui o *protossself* aos mapas interoceptivos mestres, já que as vísceras são estruturas corporais estáveis. De qualquer forma, o interessante aqui é notar como Damásio constrói a noção de subjetividade através dos mapas do corpo e a interação entre organismo e objeto. Como o autor mesmo coloca: “(...) o organismo está envolvido em relação a algum objeto, e o objeto da relação está causando uma mudança no organismo” (Damásio, 2010, pág. 145). Desse modo, como ilustrado na figura 13, o objeto não é apenas representado pelo organismo, essa representação é direcionada ao *self* (de modo a causar, como vimos, o sentimento de pertencimento), e a própria interação entre o organismo e o objeto causa mudanças nos mapas do corpo. Sendo assim, Damásio estabelece a noção de subjetividade a partir da interação entre o organismo e o ambiente a seu redor.

É importante notar que essa interação entre organismo e objeto (ambiente) não ocorre apenas com base nas reações fisiológicas decorrentes dessa interação, mas também com base nas representações (mapas) do corpo prévias a tal interação. Logo, o modelo de *self* de Damásio é baseado em dois tipos de imagens cognitivas: as imagens que representam o objeto e as imagens que representam o corpo.

A sensação de agência e de possuir uma mente seria derivado por um mecanismo comparável, ligando no tempo as atividades que pertencem a novas imagens de objetos com aquelas que definem mudanças no *protoself* no nível de mapas interceptivos, portais sensoriais e representações musculoesqueléticas (Damásio, 2010, pág. 245, tradução nossa).

E é através desses dois tipos de representação que se constrói a noção de *self*. Mas, de qualquer forma, é importante ressaltar que o *self*, na teoria de Damásio, não é apenas de um conjunto apático de representações do corpo e das relações entre corpo e objeto. Esse conjunto de representações é, acima de tudo, sentido: “os mapas do *protoself* são distintos, pois geram imagens não meramente corporais, mas também imagens *sentidas* do corpo” (Damásio, 2010, pág. 133 e 134, itálico do autor). Portanto, Damásio propõe um modelo que atribui características experienciais às representações cognitivas nas quais seu modelo se baseia.

E já que agora temos conhecimentos da teoria de Damásio sobre as emoções, sentimento e *self*, podemos adentrar sua teoria sobre a consciência.

A teoria da consciência

Finalmente, podemos fazer algumas considerações sobre a teoria de Damásio sobre a consciência (1999, 2010).

No livro de 1999, *The feeling of what happens*, Damásio apresentou as linhas-base de sua teoria da consciência, dando origem à famosa expressão *the feeling of what happens* (o sentimento daquilo que acontece), que abordaremos adiante.

Os dois níveis de consciência apresentados por Damásio (1999) são diretamente relacionados com os dois níveis “superiores” do *self* apresentados acima: a consciência central e a consciência autobiográfica (ou ampliada). A consciência central se dá, segundo o autor, quando as áreas cerebrais responsáveis por produzirem representações do corpo geram uma imagem não-verbal sobre como que o organismo é afetado pelos processos orgânicos decorrentes do contato com o objeto, dando origem à um objeto causador situado no tempo e no espaço (Damásio, 1999, pág. 168 e 169). Desse modo, Damásio (Ibidem, pág. 169 a 171) defende seis premissas em relação ao *self* central consciente (consciência central):

- Construção e exibição interna de um novo conhecimento (sentimento de conhecer);
- Nessa interação, tanto o organismo quanto o objeto são mapeados pelo cérebro, e esses mapas do objeto são denominados, pelo autor, de mapas de primeira ordem;
- Mapas do objeto causam mudanças nos mapas do organismo (como já dito acima);
- Essas mudanças podem originar outro tipo de mapa, o que o autor chama de mapa de segunda ordem, representando, portanto, a interação entre o organismo e o objeto;
- Os mapas de segunda ordem podem ser transformados em imagens, e estas imagens podem dar origem a novos mapas de primeira ordem [ou seja, elas se tornam objeto];
- As imagens mentais construídas através dessa interação, ou seja, através dos mapas de segunda ordem, são os sentimentos.

É importante notar que a interação entre organismo e objeto gera um tipo de sentimento caracterizado por um cérebro que mapeia o objeto ao mesmo tempo em que mapeia as mudanças corporais causadas pelo objeto (construindo, assim, o *self* central, como visto acima). Em outras palavras, o cérebro dá origem ao *sentimento daquilo que acontece*⁵¹. Para Damásio (1999, pág. 180), as estruturas responsáveis pelos mapas de segunda ordem são os núcleos do tronco cerebral, o hipotálamo, o prosencéfalo basal e os córtices somatossensoriais. Como se pode notar, a base neural da consciência central é basicamente subcortical.

Já a consciência ampliada é baseada no *self* autobiográfico, ou seja, um *self* que além de ter consciência do aqui e agora da consciência central também consegue se lembrar dessa experiência atual e utilizá-la em experiências futuras, mesmo que nunca cheguem a ocorrer externamente (como imaginar algo no futuro, ou simplesmente imaginar algo irreal) (Ibidem, pág. 195 e 196). Como já dito acima, a base do *self* central se constitui de mecanismos neurais que conseguem mapear o organismo, o objeto e o resultado dessa interação dando origem a novos mapas. Já o *self* autobiográfico consegue utilizar esses mapas para eventos futuros e atuais (baseado em experiências prévias), além de mapear que o organismo é o mesmo agente em todas essas circunstâncias, ou seja, ele dá origem a uma ideia de “eu” que se autorrefere (autoconsciência). Para Damásio (1999, pág. 221 e 222) a consciência ampliada, por se basear tanto em imagens presentes quanto passadas e ao mesmo tempo criar imagens sobre interações futuras, deve ser produzida por áreas de convergência, já que elas integram vários tipos de imagens. O autor propõe que essas áreas são os córtices temporais, os córtices pré-frontais (responsáveis pelos processos mnemônicos), os núcleos do tronco cerebral e a amígdala.

⁵¹ Nota importante. Para Damásio (2010), os sentimentos que acompanha toda experiência consciente são o *qualia*.

De qualquer modo, tanto a consciência central quanto a ampliada são base do sentimento daquilo que acontece, tanto no aqui e agora (consciência central) quanto no passado e no futuro (consciência ampliada). E essa é a base conceitual da teoria do autor, algo que se mantém na teoria de 2010, analisada abaixo.

Em 2010, Damásio formalmente conecta a sua teoria da consciência com a do *self* [por isso que preferimos discutir o *self* antes], mas sem descartar a teoria base da consciência e do *self* de 1999, o autor simplesmente a aprimora, tanto empiricamente [como discutido na teoria do *self* acima] quanto teoricamente [ao enfatizar a importância da dependência entre os dois conceitos]. Logo, as bases neurais da consciência, na teoria de 2010, são as mesmas para o *self* (Damásio, 2010), discutidas acima. Desse modo, podemos observar uma diferença significativa entre as bases neurais da consciência entre as obras de 1999 e 2010, representadas no quadro abaixo:

Tipos de consciência	1999	2010
<i>Consciência Central</i>	Núcleos do tronco cerebral, o hipotálamo, o prosencéfalo basal e os córtices somatossensoriais (pág. 195 e 196).	Tronco cerebral, colículo superior, córtex insular e o córtex anterior cingulado (pág. 135).
<i>Consciência Ampliada</i>	Córtices temporais, os córtices pré-frontais, os núcleos do tronco cerebral e a amígdala (pág. 221 e 222)	Córtices posteromediais (pág. 154 e 155).

Tabela 1. As diferenças das bases neurobiológicas da consciência entre a obra de 1999 e 2010. Tabela de nossa autoria.

Em relação à definição de consciência de 2010, Damásio argumenta que a “consciência é um estado da mente no qual o organismo tem conhecimento de si mesmo e do estado a sua volta” (Damásio, 2010, pág. 110, tradução nossa, *itálico do autor*). E essa é a definição geral desse fenômeno, logo, é visível que o autor implica a necessidade de haver pelo menos o *self* central para que o organismo possa ter consciência, pois, como visto anteriormente, a consciência central implica conhecer o objeto e contrapô-lo a si mesmo (o objeto é também caracterizado por sua distinção para com o organismo), e essa contraposição é também mapeada (representada). Mas a diferença em relação ao *self* da consciência ampliada é a duração deste *self*, ou seja, a permanência do *self* no espaço-tempo. Como dito acima, um *self* que consegue olhar para o passado e se projetar no futuro.

Mas de qualquer forma, o ponto central da teoria de Damásio (1999, 2010) se encontra na seguinte máxima: “os estados conscientes da mente são sentidos” (Damásio, 2010, pág. 111, tradução nossa, *itálico nosso*). Logo, nós sabemos que estamos conscientes porque *sentimos* que somos seres conscientes. Consciência é, acima de tudo, *sentir* o que está acontecendo e *sentir* um *self* em interação com o mundo através da qual esse mesmo *self* se altera, não sendo possível haver consciência sem emoção. E é importante apontar que, para Damásio (1999), distintamente de Tononi e Koch, por exemplo, apenas seres biológicos podem sentir, ou seja, apenas seres vivos podem possuir consciência, logo, Damásio indiretamente se posiciona contrariamente à postura desses autores.

Sendo assim, temos os princípios teóricos necessários para analisarmos os fundamentos da teoria de Damásio sobre a consciência.

Fundamentos da teoria de Damásio sobre a consciência

Assim como James e Panksepp, Damásio caracteriza a experiência subjetiva da consciência como uma experiência emocional (ou afetiva, no caso de Panksepp). Porém, ao contrário de James (1884), que, como visto acima, distingue processos corporais/emocionais de processos cognitivos/intelectuais, e Panksepp, que distingue representações cognitivas/condicionadas e representações afetivas/incondicionadas, Damásio não estabelece uma linha divisória entre processos e representações cognitivas e emocionais/afetivas. O autor, na verdade, atribui características subjetivas/experienciais às imagens cognitivas do corpo e às imagens resultantes da interação entre organismo e objeto. Logo, ao contrário de Panksepp, que atribui a experiência qualitativa ao afeto, Damásio atribui essa característica às imagens cognitivas na qual sua teria se baseia.

É interessante também ressaltar que, para Damásio, seria impossível existir consciência sem emoção, já que, segundo o autor, “(...) consciência e emoção não são separáveis” (Damásio, 1999, pág. 16). Isso é presente na obra de Damásio como um todo, já que, para o autor, a consciência é, acima de tudo, um sentimento, e este último é a experiência consciente de uma reação emocional prévia (mas isso abordaremos melhor no terceiro capítulo). Logo, assim como a emoção, não haveria consciência sem a interação entre organismo e objeto (já que a emoção é um programa de ação em relação a determinado estímulo), tanto que, como visto acima, a consciência se dá a partir da representação da interação entre o organismo e o objeto (mapas de segunda ordem), assim como um sentimento, que é construído a partir da percepção de uma reação emocional. O resultado da interação entre organismo e objeto é caracterizado por um sentimento. O organismo não apenas conhece o objeto, ele sente que conhece o objeto (sentimento de conhecimento).

As mudanças no protossself inauguram a criação momentânea do self central e iniciam uma cadeia de eventos. O primeiro evento na cadeia é

uma transformação no sentimento primordial que resulta em um “sentimento de conhecer o objeto”, um sentimento que diferencia o objeto de outros objetos do momento. O segundo evento na cadeia é uma consequência do sentimento de conhecer. Isso é uma geração de “saliência” para o objeto causador, um processo geralmente subsumido pelo termo *atenção*, uma extração de recursos de processamento para um objeto particular mais do que outros. O self central, então, é criado através da ligação entre o protossself modificado e o objeto que causou a modificação, um objeto que agora foi marcado sentimento e realçado pela atenção (Damásio, 2010, pág. 143 e 144, tradução nossa, *itálico do autor*).

De modo geral, a teoria da consciência de Damásio integra tanto os dois paradigmas da consciência analisados no primeiro capítulo quanto o modelo de emoção de James apresentado acima, já que o autor, mesmo considerando a importância da experiência subjetiva para analisar e teorizar sobre a consciência e processos emocionais, baseia seu modelo em imagens cognitivas, de modo que não haveria consciência sem que o cérebro do organismo represente tanto o estímulo (objeto) quanto a si mesmo (mapas do corpo). Além disso, como já analisado acima, a representação da própria interação entre o organismo (através tanto dos mapas do corpo quanto pelas mudanças fisiológicas que essa interação desencadeia) e o objeto é o que constitui os mapas de segunda ordem, que, segundo Damásio, é o que caracteriza a consciência central.

Capítulo 3

Análise dos fundamentos científico e filosóficos da teoria de

António Damásio

Nesse capítulo analisaremos importantes tópicos atuais ou que ainda continuam atuais, como a discussão sobre a possibilidade de consciência (emocional) em animais não-humanos, a crescente influência do modelo construtivista de emoção e o debate sobre consciência cortical vs consciência subcortical. Além disso, nós analisaremos alguns problemas teóricos inerentes à teoria de Damásio.

Antes de tudo, já que nosso objetivo aqui é contextualizar a teoria de Damásio no cenário neurobiológico cognitivista atual, é necessário primeiramente fazermos algumas considerações gerais sobre o modelo construtivista de emoção.

O modelo construtivista de emoção

A ideia de que emoção é um fenômeno construído é enraizada em duas ideias centrais: de que emoções não são conceitos cientificamente confiáveis, já que esses estados se referem a uma gama muito ampla de comportamentos que não possuem uma relação estrita entre si, por exemplo, ganhar uma medalha olímpica, sentir-se sem valor sem nenhuma razão aparente e variações cardiovasculares ao assistir a um filme de terror podem ser classificados como reações emocionais, de modo que a relação entre essas situações é tão embaçada que pode chegar a sugerir que qualquer coisa é uma reação emocional (Russell e Barrett, 1998, pág. 805); e que as *emoções não são processos inatos*.

Em 2006, Barrett (2006a)⁵² apresentou o que ela chamou de “paradoxo da emoção”, que se caracteriza do seguinte modo: mesmo que no convívio social/cultural seja comum classificar e observar tanto nossos próprios estados emocionais quanto estados emocionais de outras pessoas (com uma boa taxa de acerto), não há critérios científicos claros e não ambíguos que indiquem a presença de medo ou raiva ou tristeza (Barrett, 2006a, pág. 27). A solução que a autora dá a esse paradoxo é que “(...) as emoções não são dadas biologicamente [inatas, ou já instaladas biologicamente, como sugere Panksepp e outros, como visto nos capítulos anteriores], mas são construídas através do processo de categorização. Emoções existem, mas apenas como experiências” (Idem, ibidem), ou seja, segundo a autora, não há uma classificação precisa das emoções, e, logo, não há centros neurais a elas dedicados. Logo, Barrett (2006a, 2006b) se posiciona contra o pensamento de Panksepp (2000) de que as emoções básicas são espécies naturais, como apresentado no primeiro capítulo.

Como vimos, para Panksepp, há circuitos subcorticais especializados para cada tipo de emoção, logo, para o autor, medo, não é algo construído socialmente/culturalmente, mas sim algo herdado pela evolução. Entretanto, Barrett (2006b)⁵³ argumenta que não há evidências suficientes que contribuam com a hipótese de Panksepp. A autora argumenta, como dito acima, que se emoções fossem espécies naturais, nós teríamos métodos de classificar cada experiência emocional de modo a termos ferramentas que nos ajudassem a distinguir precisamente uma emoção da outra, por exemplo, um método que nos permitisse distinguir precisamente raiva de ódio e prazer sexual de alegria (Barrett, 2006b, pág. 33).

No passado, o campo [das emoções] foi moldado por questões como: “X é uma emoção?” “Quantas emoções existem?” “Qual padrão específico de eventos antecedentes, atividade neural, fisiologia e

⁵² Esse artigo se tornou muito influente na área, citado 934 vezes até no momento, segundo o Google Scholar.

⁵³ Citado 1079 vezes até o momento.

comportamento motor define cada emoção?” “Como evocamos instâncias puras de emoção, não contaminadas por influências contextuais?” e “Como regulamos e controlamos as emoções depois que elas foram desencadeadas?” Os desafios conceituais [impostos] a estes tipos de perguntas são recebidas com acusações de se negar a natureza biológica ou evolutiva da emoção ou mesmo negar a existência da emoção. A resposta emocional existe, pode ser funcional, e é muito provável que tenha sido dada pela evolução. Mas isso não significa necessariamente que raiva, tristeza e medo são categorias úteis para conduzir a ciência. (Barrett, 2006b, pág. 47, tradução nossa).

Como Russell (2003) argumenta, a classificação de expressões faciais (no caso, inferir qual é o estado emocional a partir de expressões faciais) depende vastamente do contexto sociocultural do indivíduo, o que invalida a proposta de Darwin (1872) e Ekman (1992a, 1992b) de que existem emoções e expressões faciais emocionais básicas herdadas pela evolução. Além disso, como argumenta Barrett (2006b, pág. 40), mesmo que seja comum inferir que pessoas sorrindo estejam felizes (nos testes de pareamento entre emoções e expressões faciais), essa relação entre a emoção “felicidade” e o comportamento de sorrir não é algo confiável, já que pessoas podem sorrir sem estar felizes e estar felizes sem sorrir. De modo geral, expressões faciais, como sorrisos, raramente ocorrem sem audiência. Logo, Barrett (2006b) argumenta que a posição de que emoções são espécies naturais encontra inúmeras dificuldades empíricas as quais ainda não foram superadas.

De modo geral, segundo a visão construtivista da emoção, os estados emocionais são construídos a partir de processos psicológicos básicos (Russell e Barrett, 1999; Russell, 2003; Barrett, 2006a, 2006b, 2017), ou seja, esses estados são construídos a partir do que é chamado de *afeto central* (*core affect*). Esse modelo está baseado no fato de que a maneira mais efetiva de classificar os estados emocionais, seja pela face, voz ou corpo, é através de escalas de prazer e desprazer (Barrett, 2006a, pág. 20), ou seja, através de escalas de valências (como visto no primeiro capítulo).

Essa forma de resposta afetiva é chamada de “(...) ‘central’ porque ela é influenciada por uma forma muito simples de análise de significado – se

estímulos ou eventos são úteis ou prejudiciais, recompensadores ou ameaçadores para uma determinada pessoa em um determinado momento e se uma resposta comportamental ativa é necessária” (Ibidem, pág. 31, tradução nossa).

Como vimos no primeiro capítulo, afeto é associado com valência, motivação e excitabilidade, onde valência classifica a natureza do estímulo ou situação, motivação está associada com o comportamento em direção ao estímulo ou evasão do mesmo, e excitabilidade diz respeito ao grau de intensidade do afeto (intenso ou fraco).

Mas, segundo Russell (2003), deve ficar claro que o afeto central não é o mesmo que estado emocional. Segundo o autor, o estado emocional é um estado cognitivo, pois estados emocionais possuem um objeto (tem intencionalidade). Na verdade, a visão de Russell (2003) sobre o afeto central é muito próxima da visão de Panksepp sobre afeto. Para Russell, o afeto central é um fenômeno psicológico irreduzível.

A experiência subjetiva é simples e primitiva e, portanto, irreduzível a qualquer outro fenômeno psicológico. A ciência pode buscar suas causas e consequências, mas uma análise adicional da temperatura sentida em si nos leva, não aos átomos psicológicos constituintes, mas à biologia. (Russell, 2003, pág. 148).

Para o autor, embora os processos emocionais sejam cognitivos, o afeto central, mesmo que mental, não é nem cognitivo nem reflexivo. O afeto central, ao contrário da emoção, não se direciona a nada, não possui objeto. Como dito acima, o afeto central está vinculado apenas à natureza do estímulo ou situação. O afeto central não se direciona a nada. Podemos, de acordo com Russell (2003), experienciar o afeto central sem que haja um estímulo conhecido (sem que ele se direcione a algo), como no caso das variações de humor (Idem, ibidem)⁵⁴.

⁵⁴ Aqui cabe uma nota de atualização/esclarecimento. O afeto central, segundo Barrett (que, desde 2006 vem se posicionando direta e explicitamente contra o modelo de Panksepp), na verdade, é um fenômeno *cognitivo*. Retomando a definição de cognição proposta por Neisser (1967), ou seja, que a “‘cognição’ refere-se a todos os processos pelos quais (...) a entrada sensorial é transformada, reduzida, elaborada, armazenada, recuperada e usada” (Neisser, 1967, pág. 4), Duncan e Barrett (2007) argumentam que esse tipo de definição também se refere ao afeto. A posição dos autores diz respeito ao fato de que o afeto é vastamente distribuído ao longo do cérebro, envolvendo, inclusive, áreas responsáveis por processos “cognitivos”, como os córtices pré-frontais.

Indo direto ao ponto, o modelo construtivista defende que os estados emocionais se dão tanto a partir das variações no afeto central quanto a partir da agregação de inúmeros processos cognitivos, como conhecimento incorporado e linguagem (Russell, 2003; Barrett, 2006a, 2017). Para Barrett (2006a), os estados emocionais obedecem às mesmas regras de categorização. Para ela, reconhecer o próprio estado emocional é como reconhecer uma cor (Ibidem, pág. 27). Quando nós vemos uma cor, nós diretamente associamos essa cor a um conceito linguístico, como azul ou vermelho, e isso é dado a partir de conhecimentos prévios sobre essa cor. Nós inclusive categorizamos os comportamentos de outras pessoas, atribuindo tanto um significado quanto um propósito para uma ação alheia. Desse modo, segundo esse modelo, nós aprendemos a identificar e categorizar nossos próprios estados emocionais, de modo que não temos como experienciar um estado emocional sem categorizá-lo. Por exemplo, quando nos sentimos tristes, essa emoção é construída tanto através de variações no afeto central quanto a partir de nosso conhecimento incorporado sobre a mesma, no caso, nós aprendemos a categorizar essa reação como tristeza (o que envolve tanto conhecimento incorporado quanto linguagem)⁵⁵.

De modo geral, as principais diferenças entre o modelo construtivista de emoção e o modelo tradicional de emoção (de Panksepp e Damásio, por exemplo) são as seguintes (Barrett, 2006a, pág. 39):

- Segundo a visão tradicional de emoção, há algumas emoções privilegiadas, que são geralmente chamadas de “emoções básicas”, mesmo que elas variem de teoria para teoria. Já, segundo o modelo construtivista, emoções não são espécies naturais, logo, não existem emoções básicas. Segundo o construtivismo,

⁵⁵ Aqui cabe lembrar que, para Panksepp e Solms, em Panksepp et al. (2016), em contrapartida, cognição está associado com representações condicionadas (aprendidas), logo, para o autor, conhecimento incorporado e linguagem seriam representações condicionadas que são distintas das representações afetivas incondicionadas. Para Panksepp, os afetos crus continuam existindo mesmo sem nenhuma representação condicionada.

as pessoas simplesmente transmitem (comunicam) categorias de afeto central através de conceitos que elas aprenderam;

- A visão tradicional defende que as emoções possuem mecanismos neurais específicos. Já o construtivismo defende que há dois sistemas neurais distribuídos: um que computa o valor (variações do afeto central) e outro que implementa conhecimento conceitual sobre uma emoção em particular;
- Segundo a visão tradicional, os circuitos emocionais são inatos e homólogos entre os mamíferos. Segundo o construtivismo, é o afeto central que é inato e homólogo entre os mamíferos. Segundo esse modelo, os humanos provavelmente possuem mecanismos (cognitivos) inatos de categorização, de modo que variações do afeto central possam ser categorizadas e comunicadas entre os indivíduos da espécie;
- Para a visão tradicional, ou os conceitos emocionais são inatos ou eles são diretamente derivados de categorias comportamentais já estabelecidas. Já para a visão construtivista, os estados emocionais dependem da linguagem;
- Segundo a visão tradicional, emoções como raiva, alegria, medo, etc., são mecanismos reais que modulam o comportamento. Segundo o construtivismo, ao contrário, não há ferramentas empíricas precisas para distinguir uma emoção da outra, mas é possível investigar se determinado conceito emocional é ou não normativo em determinada cultura.

É também importante notar que, como também apontado por Barrett e Satpute (2017), o modelo de LeDoux se enquadra no modelo construtivista de emoção, mesmo que, especificamente, o modelo de LeDoux tenha diferenças significativas em relação ao modelo de Barrett (2006a). Como vimos no primeiro capítulo, LeDoux e Brown (2017) propuseram um modelo de consciência emocional através do qual os autores

argumentam tanto que a emoção é um processo cognitivo e não-inato (construído) quanto que os estados emocionais são construídos a partir do que eles chamam de *self-esquemas*, através dos quais é possível agregar conhecimento autobiográfico na construção dos processos emocionais pelas GNC (redes gerais de cognição). Logo, assim como Barrett e Russel, LeDoux propõe um modelo que defende que as emoções, ao invés de inatas (inatamente construídas), são cognitivamente construídas através da integração de inúmeros processos cognitivos. Mas, de qualquer modo, o modelo de Barrett e LeDoux possuem diferenças significativas que devem ser aqui salientadas:

1) Barrett (2006b, 2017) defende que as emoções não possuem circuitos neurais específicos, sendo, na verdade, apenas uma categoria conceitual, como “dinheiro” (Barrett, 2017, pág. 13); no entanto, como vimos, LeDoux e Brown (2017) propõem que os estados emocionais são processados pelas GNC, assim como qualquer processo consciente;

2) O modelo construtivista da emoção de Barrett baseia-se no chamado “afeto central”, que se baseia em valência e excitabilidade, onde a valência é processada principalmente por mecanismos do córtex orbitofrontal, enquanto a excitabilidade é principalmente instanciada pela ativação da amígdala (Wilson-Mendenhall, Barrett e Barsalou, 2013). LeDoux (2012), no entanto, argumenta que a valência e a excitabilidade são provavelmente processos inconscientes ligados a circuitos de sobrevivência, que participam apenas indiretamente na construção de estados emocionais pelas GNC (LeDoux e Brown, 2017).

3) De acordo com o modelo de Barrett, as categorias emocionais são circunscritas por meio da modulação fisiológica do afeto central (que consiste em valência e excitabilidade) (Barrett, 2006b). Por outro lado, de acordo com o modelo de LeDoux, esses processos só participam indiretamente na construção de estados

emocionais pelas GNC. Ou seja, de acordo com LeDoux, são as GNC que tem a prioridade funcional.

Em que pese essas diferenças, tanto o modelo de LeDoux quanto o de Barrett são exemplos proeminentes do paradigma construtivista da emoção. É importante ter em mente as ideias centrais desse tipo de modelo nos tópicos que serão apresentados a seguir, já que as abordagens construtivistas e inatistas da emoção apresentam posições distintas no que diz respeito a diversos aspectos da emoção e consciência emocional, entre eles, possibilidade de consciência (emocional) em animais não-humanos.

A possibilidade de consciência animal não-humana

Além do tópico de consciência animal continuar sendo um tópico relevante em diversas áreas (como filosofia, neurociências, psicologia e ciência cognitiva), ele nos ajuda a analisar melhor a teoria de um determinado autor, pois cada teoria da consciência é a base para analisarmos a possibilidade de consciência animal de acordo a própria teoria. Por exemplo, se um determinado autor defende um modelo cortical de consciência, torna-se difícil defender a possibilidade de consciência em animais que não possuem córtex.

No entanto, ao contrário de Panksepp (1998, 2000, 2005, 2010a, 2010b, 2011; Panksepp et al., 2017), que foi um autor proeminente que dedicou grande parte de seu esforço acadêmico à questão da consciência não-humana, Damásio não se aprofundou muito sobre esse tópico, mas, mesmo assim, analisar seu posicionamento sobre essa questão nos ajuda a ter uma noção melhor sobre sua teoria da consciência.

Para Damásio (2010), a consciência “não é uma coisa” (Damásio, 2010, pág. 120), logo, segundo o autor, a consciência central e autobiográfica não são categorias rígidas, podendo haver, na verdade, inúmeros níveis entre elas, de modo que se pode

argumentar que os níveis mais baixos da consciência não são humanos. Desse modo, Damásio não toma a consciência humana como um ponto de partida para entender a natureza e o surgimento da consciência. Mas, de qualquer forma, o autor deixa claro que a consciência humana é, sem sombra de dúvidas, o tipo de consciência mais complexa que existe. “O fato de que a consciência humana, no seu alcance mais elevado, seja extremamente complexa, de grande alcance e, portanto, distinta, é tão óbvio que não requer menção” (Damásio, 2010, pág. 120, tradução nossa).

Damásio (2010) deixa bem claro que *não há como provar* satisfatoriamente se animais que não possuem linguagem têm ou não consciência, seja central ou autobiográfica, mas a autor argumenta que há como triangular determinados tipos de evidência pelos quais é possível argumentar com um bom grau de probabilidade que determinado animal possui consciência. O autor propõe dois tipos de triangulação: se a espécie apresenta comportamentos que são mais bem explicados a partir de um cérebro com mente do que comportamentos instintivos em si; e se a espécie possui os componentes *neurais básicos* descritos no capítulo anterior (Idem, ibidem).

Damásio (2010) deixa claro que a consciência central não depende de linguagem e provavelmente surgiu antes da linguagem. Segundo o autor, é bem esperado que não houvesse linguagem sem consciência central. Já a consciência autobiográfica depende extensamente da linguagem. Logo, Damásio (2010) postula que é bem provável que a consciência autobiográfica seja uma característica exclusivamente humana.

De modo geral, podemos ver que, assim como Panksepp, Damásio utiliza um modelo comportamental não-linguístico para atribuir ou não consciência a animais não-humanos, além disso, Damásio argumenta que há requisitos neurais básicos para que se possa propor que determinada espécie possui ou não consciência. Como vimos no capítulo anterior, Damásio (2010) atribui a consciência central às seguintes estruturas:

tronco cerebral, colículo superior, córtex insular, córtices somatossensoriais e córtex cingulado anterior. Porém, Damásio (2010, pág. 82) argumenta que lesões corticais, como no córtex insular e nos córtices somatossensoriais, não são suficientes para interromper a construção de estados sentimentais:

Dizer que os sentimentos surgem subcorticalmente é plausível dada a evidência (...) [de que]: danos completos aos córtices insulares na presença de estruturas intactas do tronco do cérebro são compatíveis com uma ampla gama de estados sentimentais; crianças hidroencefálicas que não possuem córtex insular e outros córtices somatossensoriais, mas que possuem estruturas intactas no tronco cerebral, exibem comportamentos sugestivos de estados sentimentais. (Damásio, 2010, pág. 92 e 93, tradução nossa).

Portanto, podemos observar que, mesmo que Damásio tenha incluído áreas corticais no seu modelo de consciência central, o autor não classifica tais áreas como fundamentais para a construção desse tipo de consciência. Desse modo, Damásio continua defendendo um modelo subcortical de consciência central.

Entretanto, o posicionamento tanto de Damásio quanto de Panksepp sofre inúmeras críticas advindas de autores que defendem modelos cognitivistas de consciência (perceptual e/ou emocional). Como dito acima, tanto Damásio quanto Panksepp propõem um modelo de consciência animal que toma comportamentos externos não-linguísticos como base para a detecção de consciência em animais não-humanos. No entanto, esse tipo de posicionamento é muito difícil de defender segundo alguns autores. LeDoux (2015), por exemplo, argumenta que não há como atestar a existência de atividade consciente em animais que não possuem linguagem, algo que, como vimos, também é argumentado por Damásio. Porém, LeDoux, ao contrário de Damásio, não estabelece princípios de triangulação para que seja possível propor que determinada espécie possui ou não consciência. LeDoux (2015) argumenta que é bem provável que animais não-humanos possuam consciência de criatura (*creature consciousness*), que é caracterizada pela responsividade a estímulos externos. Porém, o

autor argumenta que comportamentos externos não-linguísticos não são suficientes para propor que animais não-humanos possuem ou não consciência fenomênica (*mental state consciousness*), pois, mesmo que esses comportamentos atestem processos cognitivos, como memória de trabalho, atenção, metacognição e resolução de problemas, e que a consciência dependa desses processos cognitivos para existir, a mera existência desses processos cognitivos não leva à consciência (LeDoux e Brown, 2017, Box 3, material suplementar online).

Também segundo LeDoux (2015), outro problema que devemos levar em consideração ao analisar a questão da consciência animal não-humana é o antropomorfismo recorrente ao atribuir consciência a esses animais. Devido à empatia com os outros animais, principalmente com animais domésticos, é comum observar inúmeras pessoas atribuindo pensamentos e intenções intrinsecamente humanas aos seus animais de estimação. Por exemplo, é comum atribuir processamento de linguagem a cachorros e gatos, mesmo que nós sabemos (idealmente) que esses animais não possuem tal capacidade cognitiva. Logo, é possível que a atribuição de processos conscientes específicos a animais não-humanos se dá muito mais devido à projeção de intencionalidade do observador humano em relação a outros animais do que realmente uma observação neutra.

Tentando responder a este problema empírico, Paul e colegas propuseram um modelo interessante através do qual, segundo os autores, seria possível avaliar o aspecto cognitivo da experiência emocional não-humana (Paul et al., 2005). No entanto, esse modelo baseia-se em processos cognitivos que, de acordo com o modelo de LeDoux e Brown (2017), por exemplo, não dependem de consciência para ocorrer, como processos de avaliação, atenção, memória e julgamento. Os processos de avaliação podem ocorrer sem consciência, como demonstrado pela ativação de circuitos de

sobrevivência não seguidos por um estado consciente emocional, como mencionado no primeiro capítulo (LeDoux & Brown, 2017); a atenção também pode ocorrer sem consciência, é a consciência que não pode ocorrer sem atenção (Koch e Tsuchiya, 2007a); memória semântica pode ocorrer sem consciência fenomênica, ou seja, esse tipo de memória ocorre sem se dirigir a si mesmo (LeDoux, 2015); e ainda há uma discussão acalorada sobre a possibilidade de julgamento/tomada de decisão inconsciente (Libet, Gleason et al., 1983) e tomada de decisão sem atenção (Dijksterhuis et al., 2006). Portanto, embora esses autores tenham tentado fornecer uma visão cognitivista da experiência emocional em outros animais, o modelo não fornece uma justificativa convincente de como poderíamos abordar parcimoniosamente a experiência emocional em animais não-humanos, pois podemos explicar esses processos cognitivos sem abordar consciência fenomênica. Novamente, como argumentado por LeDoux e Brown (2017) na “Caixa 3”, embora a consciência dependa de memória de trabalho, atenção e metacognição, a ocorrência desses processos cognitivos não é suficiente para dar origem à consciência.

Logo, segundo o ponto de vista cognitivista de LeDoux, por exemplo, o modelo de triangulação proposto por Damásio não é o suficiente para argumentar que determinada espécie possui ou não experiência consciente. De modo geral, o grande problema de propor que determinados animais não-humanos possuem consciência é que não há como definir experimentalmente se o comportamentos exibidos por tais animais é baseado em processos conscientes ou em processos cognitivos não-conscientes. Como argumentam Koch et al. (2016, pág. 317), seres humanos também exibem inúmeros comportamentos complexos baseados em processos cognitivos não-conscientes, como detecção de significado de palavras, processos aritméticos simples e digitação rápida. Logo, Koch e colegas (2016) argumentam que não há como saber se os

comportamentos, mesmo que complexos, exibidos por inúmeras espécies animais são baseados ou não em processos conscientes.

Outra crítica do cognitivismo às teorias de Panksepp e Damásio diz respeito a quais processos neurais instanciam a consciência. Como vimos no primeiro capítulo, todos os modelos cognitivistas apresentados defendem que a consciência é um processo cortical, logo, a partir desse modelo, não haveria como propor que há atividade consciente em animais que não possuem córtex. E, segundo LeDoux (2015), propor que a consciência é um processos subcortical enfrenta problemas no que diz respeito à investigação empírica, algo que, segundo LeDoux (2015), ocorre com o modelo de consciência de primeira ordem de Block (2007).

Tanto LeDoux (2015) quanto Naccache e Dehaene (2007, pág. 518) argumentam que o modelo de consciência de primeira ordem proposto por Block é muito difícil de atestar empiricamente. Segundo o modelo de Block, a própria representação do estímulo é consciente, de modo que isso não depende da consciência de acesso (Block, 2007, pág. 484). Para o autor, a experiência consciente é inerente ao processamento de primeira ordem do estímulo, ou seja, diferentemente do modelo defendido por LeDoux e Brown, como apresentado acima, esse modelo não atribui experiência consciente à re-representação do estímulo por áreas de alta ordem. Para Block, os fenômenos periféricos são sempre conscientes. Por exemplo, objetos no nosso campo visual periférico ou uma sensação tátil periférica são sempre conscientes, mesmo que só nos atentamos a esse fato quando dirigimos a nossa atenção a eles. No entanto, como argumenta LeDoux (2015) e Naccache e Dehaene (2007), só é possível atestar tal coisa quando o estímulo periférico fica disponível à consciência de acesso, por exemplo, só podemos verificar se o participante estava ou não consciente do estímulo quando perguntamos a ele, o que necessariamente envolve linguagem. Logo, segundo LeDoux

(2015) e Naccache e Dehaene (2007), apenas a consciência de acesso (ou consciência ampliada, no modelo de Damásio) seria passível de confirmação empírica.

Desse modo, a partir desse ponto de vista, apenas a consciência humana (consciência ampliada), por ser baseada em linguagem, é possível de ser atestada empiricamente. E, já que esse tipo de consciência é vastamente instanciado por processos corticais (inclusive no modelo de Damásio, pois, como vimos no segundo capítulo, Damásio atribui a consciência ampliada a processos do córtex posteromedial), não seria possível afirmar que processos subcorticais possuem experiência consciente por si sós. Como já visto, Panksepp atribui a experiência fenomênica ao afeto, e propõe que o afeto é algo inerente aos processos subcorticais que o instanciam, logo, como afirma LeDoux (2015), a consciência afetiva no modelo de Panksepp é similar à consciência no modelo Block, já que ambos atribuem a experiências consciências a processos de primeira ordem, o que, como visto acima, é criticado por ser difícil (ou impossível) de atestar empiricamente.

Obviamente, essa breve apresentação sobre a problemática da consciência animal não-humana não encerra um tópico tão complexo como esse, mas esperamos que essa breve análise ilustre introdutoriamente as abordagens subcorticais no contexto cognitivista atual, de modo que as próprias teorias de Damásio e outros autores devem ser discutidas tendo em vista esse contexto neurobiológico atual.

A teoria da emoção, *self* e consciência de Damásio no contexto neurobiológico atual

A teoria das emoções e sentimentos

Como visto acima, o modelo construtivista de emoção não admite que o conceito de “emoção básica” seja um conceito cientificamente confiável para experimentação empírica, já que a noção de emoção depende do conhecimento autobiográfico sobre esse

fenômeno psicológico, e termos como “alegria” e “tristeza” não possuem uma categorização comportamental específica. Essa abordagem possui implicações diretas para a teoria das emoções e sentimentos de Damásio (1999, 2010).

Damásio (2010), assim como Panksepp (1998, 2000, 2005, 2007), argumenta que há circuitos neurais subcorticais dedicados à manutenção dos estados emocionais, de modo que, mesmo que haja danos generalizados à coluna espinhal, por exemplo, é possível que as emoções e sentimentos básicos ainda continuem existindo desde que haja algum tipo de conexão entre cérebro e corpo. Como o autor coloca:

“(…) enquanto que o nervo vago, que corre paralelo à medula espinhal, é preservado (que é o que geralmente ocorre), a sinalização cruzada entre cérebro e corpo permanece robusta o suficiente para garantir o controle autonômico, para operar emoções e sentimentos básicos, e para manter os aspectos da consciência que requerem *inputs* do corpo”. (Damásio, 2010, pág. 50, tradução nossa).

Logo, assim como visto acima, desde que algumas estruturas subcorticais permaneçam intactas, como o tronco cerebral superior, Damásio (2010, pág. 92 e 93) defende que as emoções e os sentimentos básicos continuam existindo. Desse modo, tanto Damásio quanto Panksepp defendem que as emoções básicas possuem circuitos subcorticais dedicados, de modo que as emoções e sentimentos básicos só deixam de existir quando essas estruturas subcorticais são lesionadas de modo generalizado.

Em resposta ao artigo de Barrett (2007b), no qual a autora argumenta que as emoções básicas não são espécies naturais, como visto acima, Panksepp (2007, pág. 281) argumenta que a teoria de Barrett está mais para uma conjectura teórica do que uma conclusão derivada de dados neurocientíficos robustos. Panksepp (Idem, ibidem) argumenta que a teoria de Barrett (2007b), na verdade, diz respeito aos aspectos secundários (aprendizado e pensamento) e terciários (pensamento sobre pensamento) dos estados emocionais, mas não aos aspectos primários (sentimento) dos estados emocionais. Segundo Panksepp, no dia-a-dia das nossas vidas emocionais é muito

difícil distinguir um aspecto do outro na construção de cada estado emocional, desse modo, segundo o autor, é possível que a teoria de Barrett (2007b) tenha levado em consideração apenas os aspectos secundários e terciários, ao invés de levar em consideração todos os aspectos da experiência emocional.

Como vimos no segundo capítulo, Damásio (1994, 1999, 2003, 2010) também distingue emoções básicas de emoções sociais, onde as emoções sociais dependem de aprendizado cultural. Logo, é possível que Damásio, assim como Panksepp, também considere que o modelo da teoria de Barrett ignora ou não compreende os aspectos básicos e primitivos da experiência emocional.

A fim de buscar falhas na teoria construtivista de Barrett e Satpute (2017), Panksepp (2007, pág. 285) argumenta que a autora deveria ter levado em conta os experimentos de estimulação elétrica cerebral localizada (EEC) e estimulação química cerebral localizada (EQC). Segundo Panksepp, essas técnicas têm demonstrado que as emoções básicas – como busca, raiva, medo, luxúria, cuidado, pânico e diversão (no modelo de Panksepp) – estão inatamente instaladas no cérebro, já que, segundo o autor, dados com EEC e EQC têm demonstrado que a estimulação de determinadas áreas cerebrais dão origem a determinados comportamentos afetivos. Panksepp (Ibidem, pág. 287) defende que os dados de EEC e EQC em animais não-humanos (ou seja, dados sobre consequências comportamentais a partir de determinada EEC e EQC) podem ser correlacionados com as características emocionais humanas, de modo que seja possível triangular as regiões cerebrais responsáveis para cada emoção básica da classe mammalia.

Em resposta a esse artigo de Panksepp, Barrett et al. (2007) argumentam que as pesquisas sobre resultados comportamentais a partir de EEC e EQC tanto em animais não-humanos quanto em humanos [tanto as citadas por Panksepp, 2007, quanto outras

pesquisas revisadas por Barrett e colegas, 2007; a maioria da década de 1960 e 1970] utilizavam técnicas com baixo grau de precisão. Ao contrário do que proposto por Panksepp, os dados revisados por Barrett e colaboradores demonstram que as emoções relatadas (no caso de participantes humanos) mudam de participante para participante mesmo estimulando as mesmas regiões cerebrais, sem que eles possuam diferenças anatômicas perceptíveis. Por exemplo, mesmo estimulando as mesmas regiões, alguns participantes relataram tristeza, e outros, felicidade. Além disso, evidências com EEC demonstram que a estimulação elétrica em animais não-humanos também dá origem a diferentes reações comportamentais, que muitas vezes estão ligadas com o comportamento comum do animal e contexto de grupo, por exemplo, a estimulação na amígdala desencadeou reações violentas apenas em animais que já possuíam histórico violento (Barrett et al., 2007, pág. 301). Desse modo, o mesmo tipo de explicação pode ser dado aos participantes humanos, pois diferentes contextos socioculturais e emocionais alteram a bioquímica neural, podendo, portanto, dar origem a diferentes reações emocionais.

Outro ponto importante levantado por Barrett et al. (2007, pág. 302) é que as estruturas anatômicas para as emoções básicas encontradas em animais não-humanos não são compatíveis com a humana. Segundo os autores, o tamanho do cérebro, a conectividade e a coordenação neural implicam diretamente no processamento emocional. Um exemplo simples disso é a significativa ativação de áreas *neocorticais* (não presente na maioria das espécies) durante experiências emocionais, como o córtex pré-frontal ventromedial e o córtex cingulado anterior, e essas áreas possuem grande conectividade com áreas subcorticais, como o hipotálamo e a PAG. Logo, segundo a teoria construtivista de Barrett, seria muito difícil de afirmar que realmente existem emoções básicas compartilhadas entre membros da mesma classe, e ainda mais difícil

entre membros de filos distintos, por exemplo, afirmar que peixes e insetos possuem consciência emocional.

Algo que poderia ser argumentado pelos defensores das bases subcorticais da emoção é que as áreas neocorticais são responsáveis apenas pelos aspectos de alta-ordem da experiência emocional. Por exemplo, seguindo a visão de Panksepp e Damásio, poder-se-ia afirmar que as áreas corticais são, na verdade, responsáveis por instanciar emoções e sentimentos sociais (que dependem de aprendizado e linguagem). O problema, segundo LeDoux (2015), como visto na seção anterior, é que a pesquisa sobre emoções depende vastamente do relato verbal do participante. Como visto no capítulo anterior, por exemplo, Damásio et al. (2000) investigaram os correlatos neurais de “emoções básicas” com base em relatos verbais sobre sentimentos autogerados, ou seja, os experimentos descritos nesse artigo foram embasados na experiência emocional autoconsciente. Desse modo, como seria possível atribuir essas mesmas experiências emocionais a animais que não possuem autoconsciência e linguagem? Logo, como que a visão de emoções básicas inatas pode se sustentar? Esses são questionamentos que LeDoux (2015) impõe aos modelos subcorticais inatista da emoções, como os modelos de Panksepp e Damásio.

Portanto, tanto segundo Barrett e colegas (Barrett, 2006a, 2006b; Barrett et al., 2007) quanto LeDoux (2015), não há evidências suficientes que corroborem a alegação de que as emoções básicas são *espécies naturais* herdadas e compartilhadas por outros animais (principalmente animais distantes de nós na cadeia evolutiva, como animais de outras classes e/ou filos). Assim sendo, é possível que a própria ideia de emoção, segundo o ponto de vista cognitivista, tenha de passar por uma revisão/reformulação conceitual, de modo que haja uma distinção entre processos fisiológicos não-emocionais, mas que constituem os processos emocionais indiretamente, e processos

emocionais conscientes construídos através de um conjunto de processos cognitivos e fisiológicos não-conscientes (LeDoux, 2012; LeDoux e Brown, 2017), como a reformulação conceitual proposta por LeDoux (2012, 2015; LeDoux e Brown, 2017), como analisada no primeiro capítulo. Por exemplo, como vimos, LeDoux e Brown (2017) revisam uma grande quantidade de estados que propõem que a ativação dos circuitos de sobrevivência (muitos deles subcorticais) não são o suficiente para dar origem a estados emocionais conscientes (sentimentos).

Esses dados realmente representam uma ameaça ao modelo teórico de Panksepp (2005; Panksepp et al., 2016), pois, como vimos no primeiro capítulo, Panksepp defende que a experiência afetiva está incorporada nos próprios mecanismos subcorticais afetivos, logo, segundo esse modelo, a ativação de estruturas ligadas à emoção básica “medo” deveria ser o suficiente para desencadear a experiência afetiva de medo. Porém, como apresentado no segundo capítulo, Damásio (1999, 2010) considera que o estado emocional consciente (sentimento) só ocorre após tanto o estímulo quanto as reações corporais serem representadas (o que inclui processos cognitivos corticais da ínsula, córtex cingulado anterior e córtex somatossensorial). Portanto, segundo o modelo teórico de Damásio, a não-representação de um estímulo mascarado implica, logicamente, que não haveria como construir o estado emocional consciente (sentimento). E em relação aos danos à amígdala, Damásio considera que áreas do tronco encefálico superior são também responsáveis pelas reações emocionais, como o núcleo parabraquial e a massa cinzenta periaquedutal (Parvizi e Damásio, 2001; Damásio e Carvalho, 2013). Assim, o modelo inconsciente de emoção de Damásio ainda permanece teoricamente válido diante desses problemas empíricos. Mas, de qualquer forma, a teoria de Damásio, segundo LeDoux (2015), continua enfrentando alguns problemas metodológicos.

Como já visto, mesmo sobre sentimentos (experiência emocional consciente) Damásio teoriza que estruturas subcorticais são também responsáveis por tais reações conscientes, juntamente com algumas áreas corticais, como a ínsula, o córtex cingulado anterior, o córtex orbitofrontal e o córtex somatossensorial. Mas, de qualquer forma, o autor explicitamente teoriza que as estruturas subcorticais seriam suficientes para levar a cabo experiências emocionais (Damásio, 2010). Um dos problemas dessa posição, como argumenta LeDoux (2015; LeDoux e Brown, 2017), é que não há como comprovar empiricamente (pelo menos não ainda) que estruturas subcorticais possam desencadear experiências conscientes por si só. Ainda mais que, como visto logo acima, a ativação de estruturas subcorticais não dá origem necessariamente à experiência emocional consciente (LeDoux e Pine, 2016; LeDoux e Brown, 2017). LeDoux (2015) argumenta que mesmo se considerássemos que estruturas subcorticais pudessem dar origem a representações de primeira ordem, só conseguiríamos atestar isso através do correlato consciente, o que necessariamente demanda estruturas corticais. Logo, como poderíamos comprovar que processos subcorticais podem ser conscientes por si só se precisamos de processos corticais para atestar isso? O que é tanto um problema teórico quanto empírico. E, como considerado no primeiro capítulo, uma vez que são as áreas corticais as responsáveis por representarem informação sensorial (Penfield e Boldrey, 1937), não haveria como existir consciência sensorial (o que inclui experiência emocional) sem representações sensoriais provenientes de estruturas corticais, logo, não haveria, segundo essa visão, como existir consciência (emocional ou de qualquer outro tipo) sem processo cortical. Desse modo, como visto acima, LeDoux (2012, 2015; LeDoux e Brown, 2017) classifica como “estado emocional” o próprio estado consciente que o caracteriza, ao invés de dividir seu modelo em “emoção inconsciente” (que só é atestada a partir do sentimento consciente subsequente) e sentimento

consciente. Assim sendo, as diferenças de classificação conceitual entre LeDoux e Damásio podem ser representadas do seguinte modo:

Diferença de classificação conceitual	<i>Damásio</i>	<i>LeDoux</i>
<i>Estado emocional</i>	Divide em duas categorias, uma consciente (sentimento) e outra inconsciente (emoção), onde o que atesta a reação inconsciente é a reação consciente subsequente, que deve ser condizente com reação inconsciente anterior.	Só um tipo de classificação de estado emocional, classificado e atestado pela própria experiência consciente que o caracteriza.

Tabela 2. Diferenças de classificação conceitual dos estados emocionais entre Damásio e LeDoux. Tabela de nossa autoria.

Tomando o exemplo da ativação dos circuitos de sobrevivência, LeDoux (2012, 2015; LeDoux e Brown, 2017) defende que os processos inconscientes decorrentes da apresentação de um estímulo potencialmente danoso representam, na verdade, a ativação do circuito de sobrevivência, e não o desencadeamento de um estado emocional em si. Isso pode ser exemplificado pela situação hipotética de ocorrer ativação dos circuitos de sobrevivência (como congelamento) diante da apresentação de um galho retorcido no chão (LeDoux, 2015), que, por ser um estímulo similar ao estímulo ameaçador inato de “cobra”, é interpretado pelo tálamo visual como um estímulo potencialmente danoso, enviando, assim, *inputs* para a amígdala, que desencadeia os circuitos de sobrevivência.



Fig. 11. Cobra ou galho? (Ilustração retirada de LeDoux, 2015).

Para LeDoux (2012, 2015), a reação comportamental de sobrevivência decorrente da apresentação do estímulo visual acima não é emocional, mas sim um mecanismo de defesa biologicamente selecionado pela evolução. Ou seja, essas reações *não são reações de medo*, medo é apenas a experiência consciente que ocorre muito depois (em escala temporal neural) de os mecanismos de defesa terem sido iniciados. Para o autor, a experiência consciente de medo só ocorre quando o estímulo for processado pelo córtex visual e então enviado às GNC (LeDoux e Brown, 2017). Como vimos no primeiro capítulo, LeDoux e Brown defendem que os estados emocionais são processos cognitivos conscientes decorrentes da integração de inúmeros processos fisiológicos e cognitivos não-conscientes pelas GNC (que se constitui principalmente de mecanismos pré-frontais).

Como apontado por Koch et al. (2016, pág. 310), um problema com as teorias subcorticais da consciência (emocional ou de qualquer tipo) é a confusão entre os pré-requisitos da consciência e o que é a própria consciência. Koch et al. (Idem, ibidem)

argumentam que é possível que mecanismos subcorticais não sejam necessários à consciência, e mesmo que sejam, é possível que eles não sejam suficientes para instanciar atividade consciente. Desse modo, seguindo essa visão, é possível que o modelo de Damásio confunda reação emocional com os pré-requisitos para que isso seja possível, classificando mecanismos inconscientes e separados no cérebro como “emoção”, ao invés de unificar o conceito entre emoção e sentimento e distinguir os mecanismos pré-requisitos da experiência emocional em si e o que a caracteriza neurobiologicamente, assim como feito por LeDoux (como pode ser visto na Tabela 2).

De todo modo, o que é importante aqui é analisar a problemática na qual a teoria das emoções e sentimentos de Damásio está inserida, a qual pode ser expandida para sua teoria do *self* e da consciência.

A Teoria do Self

O nível do *self* que mais enfrenta problemas, na teoria de Damásio, é o *self* central, tanto por motivos teóricos quanto empíricos. Como vimos no capítulo anterior, o tronco cerebral, segundo Damásio, é um dos responsáveis pela sustentação do *self* central, já que essa área é uma das grandes responsáveis pela produção dos mapas do corpo, essenciais para a constituição do *protosself*, sem o qual não é possível haver nenhum dos outros níveis do *self*, e, conseqüentemente, não seria possível haver atividade consciente. O problema, assim como levantado por LeDoux e Brown (2017, pág. E2019), é que não há como saber se essas representações subcorticais influenciam diretamente na construção do *self* central (que é necessariamente consciente). O que esses autores sugerem é que talvez as áreas subcorticais apenas produzam representações que podem ou não (não há quesito de necessidade) participar na construção da experiência consciente. Por exemplo, como já dito, a ativação de

estruturas subcorticais, como a ativação dos circuitos de sobrevivência, não necessariamente dá origem a uma experiência consciente. Segundo os autores, sem o processamento cognitivo e cortical não há como ter experiência consciente. Logo, é possível que essas representações subcorticais não influenciem diretamente a construção do *self* central. Portanto, seguindo essa visão, Damásio deveria propor um modelo que explica como que processos subcorticais podem dar suporte a um *self* consciente (*self* central).

Além disso, a nosso ver, a teoria do *self* de Damásio possui alguns problemas teóricos. Como visto no segundo capítulo, Damásio defende que a interação do organismo com o objeto dá origem à noção de perspectiva, o sentimento de conhecer e a sensação de propriedade. Porém, o autor não explica como isso ocorre. Ou melhor, Damásio não explica tanto como e porque as estruturas cerebrais do *self* central (tronco cerebral, insula, córtex cingulado e córtex somatossensorial) dão origem a essas propriedades quanto como a interação entre o organismo e o objeto dá origem ao sentimento de conhecer e o sentimento de propriedade. Na verdade, Damásio (2010, pág. 146) apenas propõe essas relações teóricas, sem as formalizar e justificar. Desse modo, infelizmente, o autor não fornece nem um modelo conceitual nem um modelo empírico que explique, satisfatoriamente, o surgimento dessas propriedades.

Teoria da Consciência

Nessa primeira parte, avaliaremos a teoria da consciência de Damásio no contexto cognitivista atual, tentando ir além do que já foi discutido na seção “teoria das emoções e sentimentos” (já que sentimento, segundo Damásio, é um processo consciente). E, na próxima seção, faremos uma análise conceitual da teoria da

consciência e emoções de Damásio (sobre a integridade conceitual de seu modelo teórico).

Para começar, a grande distinção entre a teoria de Damásio em relação às teorias mais atuais (e *mainstream*, como no caso da GWT) diz respeito às bases neurais da consciência. Damásio, como visto acima, é um dos poucos autores de grande renome que defende, ainda hoje, que a consciência, pelo menos em sua forma base (a consciência central) possui bases subcorticais, como o tronco cerebral⁵⁶. E já que os sentimentos são a base de sua teoria sobre a consciência, o mesmo considera que tais fenômenos também são primordialmente processados por tais áreas (Damásio, Damásio e Tranel, 2013; Damásio e Carvalho, 2013).

Bem, como vimos no primeiro capítulo, os grandes proponentes da GWT e GWN, Baars e Dehaene, defendem que os córtices pré-frontais e parietais são os principais responsáveis pela atividade cortical. Podemos inclusive dizer que tais áreas corticais são o próprio espaço de trabalho global (GW), ou, em outras palavras, onde a informação é integrada e realçada pela atenção ficando acessível através da memória de trabalho. Já Tononi e Koch, a partir da IIT, defendem que a consciência é a própria integração de informação que ocorre, primariamente, nos córtices posteriores. Logo, podemos observar que, desde as teorias de Crick e Koch (1990), o corticocentrismo continua sendo o paradigma principal nas teorias neurobiológicas da consciência, com poucas exceções, como no caso de Damásio e Panksepp. Além disso, tanto a G(N)WT

⁵⁶ Aqui cabe um esclarecimento sobre a então famosa Declaração de Cambridge sobre a Consciência Animal. Pelo que parece, essa declaração é muito mais ética do que conceitual. Tanto que um de seus responsáveis é o próprio Koch, que, como vimos, é completamente contra a possibilidade de a consciência ser idêntica ao processamento de áreas subcorticais. No entanto, isso não implica que animais que não possuem córtex, ou, talvez, nem consciência, possam ser tratados de forma antiética ou cruel. A experimentação animal não deve causar sofrimento ou estresse desnecessário a qualquer animal. Isso é mais uma posição ética do que propriamente conceitual (sobre quais animais podem possuir ou não consciência), tanto que a declaração é escrita em termos de “possibilidade de” (por exemplo, é possível que polvos tenham consciência) e não afirmações categóricas (por exemplo, há evidências que comprovam que...).

quanto a IIT são teorias computacionais da consciência, pois avaliam o funcionamento neural que dá origem à consciência pelos moldes computacionais, utilizando-se do próprio vocabulário computacionalista, como buffer, codificação, decodificação, retenção, armazenamento, etc., ou seja, são teorias claramente baseadas na metáfora computacional da consciência.

Esse modelo cognitivista-computacional é o mais aceito atualmente (LeDoux, 2015), desse modo, é oportuno analisar, mesmo que brevemente, o modelo de Damásio frente a esse paradigma dominante. E, como visto acima, é muito difícil, atualmente, defender um modelo subcortical de consciência, já que seria necessário provar que tais representações subcorticais (ou de primeira ordem) são conscientes por si sós, algo que, por exemplo, não se mostrou possível para com o córtex visual (Crick e Koch, 1995), como já mencionado no primeiro capítulo. Assim, como anteriormente questionado, como distinguir a consciência do que a caracteriza? Ou seja, se só conseguimos estudar a consciência através de relatos conscientes que demandam processos de alta ordem, como provar que há representações de primeira ordem conscientes antes mesmo da experiência consciente de alta ordem? (LeDoux, 2015).

Sobre isso, é também importante lembrar-nos da *Review* de Koch et al. (2016), também discutida no primeiro capítulo, onde os autores argumentam que há poucas evidências que corroboram as hipóteses subcorticais da consciência, uma vez que lesões nessas áreas não necessariamente impedem a atividade consciente, e mesmo quando isso ocorre, a explicação mais sustentável é que isso impede, na verdade, processamentos corticais, uma vez que lesões corticais também levam ao coma mesmo quando as áreas subcorticais estão intactas.

Portanto, podemos dizer que, no que diz respeito às bases neurobiológicas da consciência, a teoria de Damásio enfrenta tanto problemas de aceitação, devido às

teorias mais em voga hoje em dia, quanto problemas metodológicos, já que não há como provar, pelo menos no estágio neurocientífico atual, que representações subcorticais são conscientes por si sós. Pois, mesmo que Damásio (2010, pág. 155) estabeleça que os córtices posteromediais sejam regiões de grande importância para o *self* autobiográfico e consciência ampliada, sendo de certo modo compatível com as atuais ideias de Koch e Tononi (Koch et al., 2016; Tononi et al., 2016), não podemos perder de vista que Damásio não considera a consciência ampliada como algo distinto da consciência central, mas sim uma extensão da mesma. Logo, o papel dos córtices posteromediais seria o de integrar informação de inúmeras áreas neurais relacionadas com a construção da consciência central com informações advindas de áreas neurais relacionadas com processos mnemônicos, como o córtex temporal e o córtex entorrinal. Ou seja, Damásio não deixa de defender que áreas subcorticais são áreas de grande importância para a atividade consciente.

Sumarizando o que foi analisado no primeiro capítulo, Crick e Koch (1990, 1993, 1995) defenderam que não existiria atividade consciente sem que representações sensoriais fossem enviadas aos córtices pré-frontais (como o exemplo da consciência visual, que só ocorre quando o processamento de um estímulo visual realizado pelo córtex visual é então enviado aos córtices pré-frontais). Tanto Baars (1988, 1998, 2002) quanto Dehaene e colaboradores (Dehaene et al., 1998; Dehaene e Naccache, 2001; Dehaene e Changeux, 2011; Dehaene, 2014) defendem que a informação só se torna consciente quando esta alcança o espaço de trabalho global, que, a partir dos trabalhos de Dehaene e colaboradores, passou a ser identificado com os mecanismos do córtex pré-frontal e parietal, formando uma rede de conexões corticais de longa distância ao longo do cérebro. Koch e Tononi (Tononi e Koch, 2015; Koch et al., 2016; Tononi et al., 2016) defendem que a consciência é um tipo de integração de informação que ocorre

nos córtices posteriores e que não necessariamente depende de processamento subcortical. E LeDoux e Brown (2017) têm proposto que a consciência emocional não é um tipo especial de consciência, sendo processada pelas GNC (formado por mecanismos pré-frontais, parietais e insulares) assim como qualquer outro processo consciente.

Apesar dessas críticas, não podemos deixar de ressaltar a perspicácia de Damásio ao vincular o sentimento à consciência. O autor observou a grande importância do “sentir algo” enquanto estamos conscientes, tanto que, no livro de 1999, *The feeling of what happens*, Damásio pede ao leitor que tente desvincular “estar consciente” de “estar sentindo”, algo que aparenta ser impossível, pelo menos no que se refere à consciência humana. Infelizmente, essa posição de Damásio é ignorada pela corrente *mainstream* no estudo neurobiológico da consciência. Mesmo Damásio sendo um neurocientista cognitivista é visível a distinção entre suas posições e a de outros autores cognitivistas. O autor, por exemplo, não defende nenhum modelo computacional de consciência e a maioria de suas metáforas conceituais diz respeito à experiência humana (inclusive artística) ao invés de algum tipo de processamento sem consciência.

Entretanto, mesmo diante de inúmeros méritos, a teoria de Damásio, a partir de um ponto de vista filosófico, enfrenta sérios problemas conceituais.

Os problemas conceituais da teoria das emoções e consciência de António Damásio

Nessa seção, faremos uma análise conceitual do modelo teórico proposto por Damásio (1999, 2010). Essa análise é de grande importância, pois nos ajudará a entender a interação entre os conceitos propostos por Damásio. De modo geral, nosso objetivo central aqui é demonstrar como o modelo conceitual de Damásio dá origem a

uma teia de problemas teóricos, visto não possuir um rigor típico de grandes obras filosóficas e/ou científicas. Trataremos apenas da base conceitual da teoria da consciência, ou seja, dos conceitos de emoção, sentimento e consciência.

Como visto no capítulo anterior, tanto no livro de 1999, *The feeling of what happens*, quanto no livro de 2010, *The self comes to mind*, Damásio aborda primeiramente o objeto das emoções e sentimentos ao invés de se dirigir diretamente ao tema da consciência, já que, não seria possível abordar a consciência sem primeiro considerar esses fenômenos profundamente, uma vez que não seria possível, segundo o autor, ter atividade consciente sem processos emocionais. Assim, antes de considerarmos o conceito de consciência, temos de retomar as definições propostas por Damásio para os conceitos de emoção e sentimento abordados pelo autor no livro de 2010, uma vez que esses conceitos passam por algumas atualizações do livro de 1999 (como visto no segundo capítulo).

De acordo com Damásio (2010), a emoção é um programa de ação amplamente automatizado o qual modula nossa expressão facial, postura corporal e funcionamento fisiológico (2010, pág. 76). Os sentimentos, por outro lado, são percepções do que acontece durante um estado emocional (2010, pág.81). A partir desse livro, como vimos no capítulo anterior, os sentimentos são sempre considerados como uma experiência consciente de um estado emocional⁵⁷. É importante notar que Damásio não define sentimentos de nenhuma outra maneira. *Não há definições secundárias para o conceito de sentimento* além do referido acima, e é relacionado a ele que, como veremos a seguir, Damásio define a consciência.

⁵⁷ Damasio não deixa claro se os sentimentos primordiais são conscientes ou não. De acordo com o autor, os sentimentos primordiais são “(...) reflexões espontâneas do estado do corpo vivo” que (...) “resultam de nada além do corpo vivo e precedem qualquer interação entre a regulação da maquinaria da vida e qualquer objeto” (Damásio, 2010, pp. 101). Então, já que eles resultam da (sensação) do corpo vivo e são a base de todos os sentimentos emocionais, é possível que estejam em um nível subconsciente ou pré-consciente da mente.

Em 2010, Damásio propôs que a consciência é “*um estado de espírito em que há conhecimento da própria existência e da existência dos arredores*” (Damásio, pág. 157, tradução nossa, itálico original). Há aqui um complemento da definição de 1999, “o sentimento do que acontece”, já que a interação entre o organismo e o objeto é sempre acompanhada por um sentimento que nos diz que o objeto (arredores) nos pertence (se refere ao *self*) por uma duração específica (enquanto estamos conscientes do objeto) (Idem, *ibidem*). Segundo o autor, nós não só conhecemos a existência do objeto e de nós mesmos, nós a sentimos. De acordo com Damásio (2010), todo estado consciente da mente é acompanhado por esse sentimento. Logo, aqui vamos manter o conceito de 1999 (que também é usado pelo autor no livro de 2010, pág. 130).

Como já comentado no capítulo anterior, Damásio recorrentemente faz uso de um tratamento poético/literário para abordar seus conceitos. Nesses livros, Damásio usa uma abordagem artística ou paisagística para discorrer sobre a questão da emoção e da consciência⁵⁸, de modo que é possível que essas definições também sejam influenciadas pela experiência estética do autor, algo que, por exemplo, é problematizado por Popper (1959/2002, pág. 22), uma vez que este autor argumenta que experiências subjetivas não podem justificar declarações científicas. Mas o nosso objetivo não é tratar do aspecto poético da teoria da consciência de Damásio, mas sim analisar o seu *background* técnico (no que diz respeito ao rigor conceitual) e científico de seu modelo teórico a partir da análise da integridade conceitual. Aqui não analisaremos como Damásio justifica seu modelo conceitual dentro de seu próprio modelo teórico de consciência, mas sim analisaremos as implicações lógicas através da interação direta entre os principais conceitos de seu modelo, neste caso, os conceitos de sentimento, emoção e consciência. Vejamos como esses conceitos interagem:

⁵⁸ Por exemplo, para Damásio (2010), “*todos os sentimentos de emoção são variações musicais complexas sobre sentimentos primordiais*” (*Ibidem*, pág. 14 e 15, tradução nossa, itálico nosso).

Definições de Damásio:

- O sentimento é a percepção do que acontece durante um estado emocional (Damásio, 2010, pág. 81) (O autor não apresenta outra definição de sentimento);
- Consciência é o sentimento do que acontece⁵⁹ (Damásio, 1999, pág. 8).

Assim, uma vez que o conceito de consciência depende do conceito de sentimento (como analisado no segundo capítulo), é possível estabelecer uma interação direta entre as definições acima:

- A consciência é uma percepção de um sentimento (sensação de percepção)⁶⁰; ou uma percepção emocional de uma percepção emocional (percepção de percepção); ou sentimento de um sentimento.

Consideremos um pouco mais detalhadamente os itens destacados acima. No modelo conceitual de Damásio, a consciência é abordada como um “sentimento do que acontece”. Tomemos apenas a última parte da frase: o que acontece. Bem, como sabemos o que está acontecendo? Quais são os mecanismos necessários para tal coisa? Provavelmente, a resposta mais simples para essa questão é: através de mecanismos sensoriais e perceptivos. Assim, a própria definição de consciência está ligada à percepção. Podemos simplesmente substituir “o que acontece” por “o que percebemos” que temos que consciência é “o sentimento da percepção”, que já foi inferido na análise conceitual acima. E uma vez que o sentimento é em si um tipo de percepção (neste caso, uma percepção de um estado emocional), podemos inferir que a consciência é “percepção da percepção”.

A percepção, de acordo com Pomerantz (2008), é o processo pelo qual é possível integrar e interpretar informação sensorial. Assim, uma “percepção de uma percepção”

⁵⁹ O que já inclui os conceitos de “self no ato de conhecer” e o “sentimento de conhecer”.

⁶⁰ Uma vez que a consciência é um sentimento e o sentimento é a percepção de um estado emocional, e a percepção de um estado emocional é um sentimento, podemos estabelecer essas implicações através de inferência conceitual simples (*modus ponens*).

requer dois níveis de integração e interpretação de informação sensorial. Pode-se dizer que, de certa forma, isso faz sentido na teoria de Damásio, uma vez que para ter consciência é necessário um tipo de mapa que mapeie a interação entre o organismo e o objeto: os mapas de segunda ordem, que são o resultado da interação de dois tipos de mapas de primeira ordem (o organismo e o objeto), cada um exigindo um tipo de processamento perceptual (mapas neurais). Em outras palavras, ao mesmo tempo em que percebemos o objeto, percebemos nosso *self* no ato de perceber (percepção da percepção). Então, de uma forma ou de outra, para ter consciência de alguma coisa, precisamos de três processos perceptivos: a percepção do corpo, a percepção do objeto, e a percepção dos resultados da interação entre o objeto e o corpo (o que inclui tanto o corpo físico quanto os mapas do corpo no *protossself*). Ou seja, precisamos do processamento perceptual do corpo e do objeto para dar origem a informações sensoriais que devem ser integradas e interpretadas por outro processamento perceptual. Em outras palavras, nós precisamos que um processamento perceptual torne-se um estímulo (informação sensorial).

Assim, quando estamos cientes de um objeto, isso acontece porque o processamento perceptual do objeto (informação sensorial) interage com o processamento perceptual do corpo através de mecanismos do tronco cerebral, córtex insular e córtex somatossensorial, dando origem a informações neuronais sensoriais (já que elas se originam dentro do cérebro), que é então processada por mecanismos perceptuais que dão origem a um segundo tipo de processamento perceptual, o que é chamado de mapa de segunda ordem. Aqui vale a pena ressaltar que essa explicação complicada é inerente ao modelo conceitual de Damásio, já que o autor utiliza os conceitos de “sentimento”, “percepção”, “emoção” e “consciência” sem se atentar nas consequências conceituais resultantes da interação direta desses mesmos conceitos.

Portanto, apenas isso, como é evidente, seria suficiente para questionar tanto a parcimônia quanto o rigor conceitual da teoria de Damásio (1999, 2010), mas os problemas não param por aqui.

Outra implicação é que a consciência é um “sentimento de um sentimento”. Como mencionado acima, Damasio (1999, pág. 69) aborda essa expressão como sinônimo de “conhecer um sentimento”. “Conhecer”, no modelo de consciência de Damásio, é também um tipo de sentimento: o sentimento de conhecer. O conceito de sentimento de conhecer está diretamente ligado ao conceito de “o sentimento do que acontece”, uma vez que o sentimento de conhecer é “o sentimento do que acontece quando um organismo está envolvido com o processamento de um objeto” (Ibidem, pág. 28), isto é, quando o cérebro do organismo cria mapas de segunda ordem, o que é o mesmo que a consciência central. Portanto, “conhecer um sentimento”, “sentir um sentimento” e consciência (central) são a mesma coisa. E, uma vez que a consciência é “sentir um sentimento” (sentimento de um sentimento), ao tomar o conceito de sentimento visto acima, temos que consciência é uma percepção do corpo durante um estado emocional causada por uma percepção anterior do corpo durante um estado emocional, isto é, a percepção do corpo durante um estado emocional deve causar outra reação emocional, cuja percepção é a consciência. Enfatizamos que essa implicação ocorre porque Damásio não fornece outro conceito de sentimento ao longo de sua obra. E outro problema teórico dessa implicação é que o sentimento (pelo menos desde o livro de 2010) é sempre uma experiência consciente, logo, ao mesmo tempo em que sentimento define consciência, consciência define sentimento, isto é, a consciência é o sentimento daquilo que acontece (ou sentimento de um sentimento) e o sentimento é uma experiência consciente de um estado emocional: é uma definição circular. Este problema teórico também ocorre com a análise acima sobre a percepção da percepção,

já que, uma vez que o sentimento é um tipo de percepção, ao mesmo tempo em que o sentimento é uma experiência consciente, esse tipo de percepção é sempre consciente. Portanto, a consciência é uma percepção emocional consciente de uma percepção emocional consciente, causando outra explicação circular, o que prejudica a replicabilidade do modelo (Schaller, 2016).

Outra implicação importante é que, uma vez que o conceito de consciência está ligado ao conceito de emoção, a consciência ocorre exclusivamente através de uma reação emocional *para qualquer tipo de experiência*. Esta implicação também é reconhecida por Damásio, uma vez que o autor afirma que “nenhum conjunto de imagens conscientes de qualquer tipo e sobre qualquer tema nunca deixa de ser acompanhado por um coro obediente de emoções e consequentes sentimentos” (Damásio, 2010, pág. 182). Para analisar isso, vamos utilizar a experiência consciente de vermelho como exemplo. Seguindo a teoria de Damásio, essa experiência é, acima de tudo, uma experiência emocional. De acordo com a teoria de 2010, podemos dizer que quando vemos um objeto vermelho, essa informação sensorial muda o status original (o corpo em si) dos mapas dos portais sensoriais direcionados externamente, além de mudar os mapas mestres do organismo, já que o organismo precisa mudar sua postura para capturar o objeto e, seguindo a teoria, ele também deve, de alguma forma, causar mudanças nos mapas mestres interoceptivos, ou seja, a experiência do vermelho causa mudanças no estado fisiológico dos órgãos internos. *A experiência perceptiva do vermelho é uma experiência emocional*. E, como discutido acima, o estímulo vermelho deve causar uma reação emocional no corpo cujas mudanças são percebidas na forma de um sentimento que, por sua vez, deve causar outra reação emocional cuja percepção é, em última instância, a experiência consciente do vermelho. Como o leitor pode ver,

seguindo as implicações conceituais acima, toda experiência consciente deve ocorrer *através do corpo duas vezes*.

Damásio (1999, pág. 8) afirma que o corpo é o teatro da emoção, o que é enfatizado no livro de 2010. Logo, uma vez que o sentimento é a percepção do corpo durante um estado emocional e a consciência é um sentimento (de um sentimento), então toda experiência consciente também tem o corpo como seu teatro. E aqui não estamos apenas considerando a cognição incorporada. Dizer que qualquer experiência consciente acontece através da reação emocional no corpo significa que não há gatilho cognitivo direto para qualquer atividade consciente que seja. Ou seja, embora a teoria de Damásio admita que essas reações também ocorram através de imagens corporais (uma vez que as mudanças físicas no corpo são só sentidas através da mediação dessas imagens), essas mudanças também devem ocorrer no corpo físico, de modo que um gatilho cognitivo direto não é teoricamente aceito pela teoria da consciência de Damásio. Assim, mesmo quando acreditamos que temos uma experiência consciente que, *teoricamente*, não exige uma reação emocional, como a experiência de vermelho, ela também requer mudanças corporais devido a uma reação emocional a ser conscientemente experienciada. Vamos dar outro exemplo. Quando alguém lhe diz que um ente querido faleceu, o cérebro, primeiramente, deve fazer uma avaliação cognitiva da afirmação, o que requer vários processos cognitivos, como atenção, memória, processamento de linguagem e assim por diante. Mas de acordo com a teoria de Damásio, nenhum desses processos torna-se consciente antes das mudanças no corpo (tanto nos mapas do corpo quanto no corpo físico), ou seja, não há gatilho cognitivo direto, o corpo deve reagir a essa informação antes que nós tomemos consciência do que foi dito e da emoção que isso causou. Logo, isso também aconteceria pelo que estamos pensando e imaginando agora. A experiência consciente das palavras que você está

lendo agora é dada por este mesmo mecanismo (através da reação emocional nos mapas do corpo e no corpo físico).

Como pode ser observado, os conceitos e definições utilizados por Damásio (1999, 2010) dão origem a uma rede de problemas teóricos, onde um conceito define o outro ao mesmo tempo em que este define o primeiro, dando origem a explicações circulares, além de implicações teórico-conceituais que estão longe de ser parcimoniosas e rigorosas. E essa falta de rigor conceitual também dá origem a outros problemas filosóficos, como a posição contraditória de Damásio (2010) envolvendo aspectos que envolvem a relação mente-cérebro.

Reduccionismo e poder causal da consciência em Damásio

No livro de 2010, *The Self comes to mind*, Damásio, pela primeira vez, aborda a questão da equivalência (identidade) entre cérebro e mente (Damásio, 2010, pág. 224) – e aqui deve ficar claro que Damásio (Idem, ibidem) considera os conceitos de equivalência e identidade como sinônimos. Damásio (Idem, ibidem) argumenta que o cérebro, assim como qualquer estrutura física, pertence ao mundo físico, onde as relações de identidade são geralmente estabelecidas através de propriedades físicas, como massa, movimento e força. Segundo o autor, aqueles que rejeitam a identidade entre mente e cérebro partem do pressuposto que os processos mentais não podem ser discutidos em termos físicos (questão de linguagem)⁶¹. Segundo o autor, “(...) o motivo dado é que até agora a ciência não conseguiu determinar os atributos físicos dos padrões mentais, e se a ciência não pode fazê-lo, então o mental não pode ser identificado com o

⁶¹ Essa postura é, de fato, adotada por alguns filósofos proeminentes como Fodor (1974) e Davidson (1995), segundos os quais, mesmo que os processos psicológicos sejam instanciados por processos físicos, o vocabulário psicológico não se reduz ao vocabulário da física (identidade de ocorrência – *token identity*).

físico. Receio, no entanto, que esse raciocínio não proceda” (Damásio, 2010, pág. 224, tradução nossa).

Segundo Damásio, nós identificamos objetos no mundo externo através do nosso aparato sensorial, pelos quais nós estabelecemos os procedimentos de medição. Entretanto, como aponta o autor, nós não podemos fazer o mesmo com os processos mentais. Mas, isso não se deve ao fato de que os processos mentais não possuam equivalentes neurais, mas sim porque eles ocorrem dentro de nossas cabeças, ficando, assim, fora do alcance das técnicas de medição externas (Damásio, 2010, pág. 224 e 225). Porém, segundo Damásio, isso não é o suficiente para defender que há uma falta de fisicalidade nos estados mentais, já que tal visão (de que estados mentais não são equivalentes a processos neurais) é embasada simplesmente na introspecção. Segundo o autor, mesmo que seja um obstáculo metodológico o fato de as imagens mentais estarem inseridas no cérebro, não haveria outro lugar no qual elas poderiam estar, o estranho (e anatomicamente impossível) seria observá-las externamente (Ibidem, pág. 225). De qualquer forma, Damásio deixa claro que o atual estágio do desenvolvimento científico e tecnológico não permite que estabeleçamos identidades precisas entre determinado processo mental e determinados processos neurais, e que, no atual momento, só nos resta realizar aproximações teóricas graduais a partir de novos dados neurobiológicos (Idem, ibidem).

O problema dessa passagem do texto de Damásio começa quando o autor aborda a questão da causação mental. De fato, se adotarmos uma posição reducionista entre os estados mentais e os processos neurais, não haveria problema ao se dizer que um determinado estado mental influencia a ação (através do controle de nossos músculos e nervos), já que, nesse caso, seria apenas um estado biológico (físico) causando outro estado biológico (físico). Porém, a justificativa que Damásio dá para o controle mental

do comportamento é um tanto quanto complicada. Segundo o autor: “uma vez que os estados mentais e os estados neurais são considerados como *as duas faces do mesmo processo*, (...), a causalidade descendente é um problema a menos” (Damásio, 2010, pág. 226, *itálico nosso*).

De certo modo, essa fala aparenta estar inserida no monismo de duplo-aspecto, segundo o qual, a interação entre mente (e consciência) e o cérebro pode ser vista como um único processo contendo dois lados, onde um não se reduz ao outro (Velmans, 2009, pág. 32). E é interessante ressaltar que essa posição é compartilhada por outros autores, como Velmans (1991) e Chalmers (1996). E, como visto no primeiro capítulo, Panksepp (2005) utiliza essa teoria para justificar sua posição teórica de que o afeto é um aspecto inerente ao funcionamento cerebral, de modo que ele não se reduz a processos cognitivos e neurobiológicos (segundo Panksepp, o afeto é o elemento qualitativo da experiência). O ponto aqui é que pelo menos Panksepp manteve o axioma central dessa teoria: de que os aspectos não reduzem um ao outro. Já Damásio, como visto acima, argumenta em prol da identidade mente-cérebro ao mesmo tempo em que declara que a causação mental pode ser teoricamente resolvida a partir da visão de que os aspectos físicos (biológicos) e mentais são duas faces (dois aspectos) do mesmo processo, onde, portanto, uma posição teórica contradiz a outra.

Essa postura teórica (de considerar que o processo neural tem dois aspectos) fica ainda mais complicada quando Damásio (2010, pág. 226) utiliza o exemplo da anatomia da mão para defender seu ponto de vista reducionista entre a atividade neural e produção de imagens mentais. Segundo o autor, a mão é feita de ossos, músculos, tendões, vasos sanguíneos e nervos (propriedades físicas), e quando esse objeto opera uma ação (fenômeno físico), não há como distinguir a mão da ação que ela executa. A nosso ver, esse exemplo de Damásio não representa muito bem sua visão de que os

processos neurais possuem dois tipos de aspectos (ao mesmo tempo em que o autor argumenta que os processos mentais e neurais são equivalentes/idênticos), já que toda ação é inerentemente física e externamente observável, ao contrário dos estados mentais, que, mesmo que físicos (a partir de um ponto de vista reducionista), não são externamente observáveis. E, de qualquer forma, ação não é um aspecto qualitativo do objeto. Na verdade, esse exemplo de Damásio seria um bom exemplo de correlação físico-físico (um processo físico correlacionado com outro).

Portanto, se Damásio continuasse defendendo uma postura reducionista restrita entre os processos mentais e neurais, ao invés de declarar que a causação mental é explicada devido ao fato de os processos neurais possuírem um aspecto físico e um aspecto mental, os problemas teóricos apresentados acima não ocorreriam. De qualquer modo, aqui cabe ressaltar que, de nenhuma forma, esses problemas teóricos e conceituais analisados nas duas últimas seções diminuem ou ofuscam a grande importância da teoria da consciência e emoções de Damásio tanto para as ciências da mente quanto para a filosofia.

Conclusão

Tomando a obra de António Damásio como fio condutor, o objetivo central dessa dissertação foi analisar a questão da emoção, sentimento, *self* e consciência desse autor no contexto neurobiológico geral, comparando a rica obra de Damásio com outros autores proeminentes, como Panksepp (que defende um modelo afetivista de consciência e emoção), LeDoux, Barrett, Baars, Dehaene, Koch e Tononi (estes últimos defendem modelos cognitivistas de consciência e emoção).

No primeiro capítulo, buscamos analisar, de modo geral, importantes teorias dos dois grandes grupos teóricos sobre a consciência nas neurociências, no caso, o paradigma afetivista, representado, principalmente, por Panksepp, e o paradigma cognitivista, que, como vimos, detém a maioria das teorias sobre a consciência atualmente (dentro do contexto neurobiológico). Ambos os paradigmas influenciam a teoria da consciência de Damásio (1999, 2010), já que, mesmo que o autor considere que a consciência é caracterizada como uma experiência emocional (posição próxima da defendida por Panksepp), o autor embasa todo seu modelo em imagens e mapas cognitivos (que Panksepp e Solms, em Panksepp et al., 2016, definem como representações condicionadas).

Já no segundo capítulo, procuramos enfatizar que, mesmo que Damásio (1999, 2010) defenda um modelo que, de certo modo, tenha relação com alguns axiomas centrais dos paradigmas afetivista e cognitivista, a teoria do autor possui diferenças pontuais e importantes em relação aos mesmos. Por exemplo, mesmo que Damásio considere que a consciência possui natureza emocional, o autor atribui a experiência qualitativa ao sentimento (consciente), diferentemente de Panksepp (2005), que atribui a experiência qualitativa ao afeto embutido na atividade neural; e, mesmo que Damásio embase sua teoria em imagens e mapas cognitivos, o autor não defende um modelo estritamente cortical e computacional de consciência, pelo contrário, Damásio (1999, 2010), como visto no segundo e terceiro capítulo, chega a considerar que o funcionamento saudável do tronco cerebral superior é o suficiente para manter os sentimentos básicos.

No terceiro capítulo, nosso objetivo foi realizar uma análise geral da obra de Damásio, analisando a teoria do autor sobre emoção, *self* e consciência frente ao modelo neurobiológico atual. De modo geral, como visto acima, os principais problemas da

teoria de Damásio (*segundo o paradigma cognitivista*) se referem tanto ao fato de que Damásio defende um modelo subcortical e inatista de emoção, que é vastamente criticado pelo modelo construtivista, quanto ao fato de que a teoria de Damásio admite que a consciência central pode ser instanciada por processos subcorticais. Além disso, analisamos como que o modelo conceitual de Damásio dá origem a problemas teóricos e filosóficos, como explicações circulares e antieconômicas e defesas de posições inerentemente contrárias (como teoria da identidade e teoria de duplo-aspecto).

De qualquer modo, esperamos que essa dissertação tenha contribuído com o debate sobre emoção, consciência, *self* e consciência tanto na filosofia da mente quanto na filosofia das neurociências (principalmente na neurofilosofia), como o debate sobre cognição incorporada, já que Damásio argumenta que emoção, sentimento, *self* e consciência não existiriam sem corpo. Além disso, como vimos, a teoria de Damásio nos oferece ferramentas teóricas para analisarmos a possibilidade de consciência (emocional) em animais não-humanos, onde o confronto entre teorias corticais e subcorticais possui uma grande relevância, já que cada modelo implica diretamente em como abordamos a possibilidade de consciência animal não-humana. E especial atenção deve ser dada à importância da experiência fenomênica na obra de Damásio, já que o autor caracteriza a consciência como uma forma de sentimento, se distanciando, assim, de teorias funcionalistas (no sentido clássico da palavra) e teorias computacionalistas (como a teoria de Tononi e Koch, que admite a possibilidade de consciência sem corpo e emoção).

De forma geral, esperamos ter humildemente contribuído com um debate que vai muito além do que foi apresentado aqui.

Referências

Anderson, A.K., & Phelps, E.A. Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. **J. Cogn. Neurosci.**, v. 14, p. 709–720. 2002.

Armstrong, D.M. **The Mind-Body Problem: An Opinionated Introduction**. Boulder: Westview Press, 1999.

Aru, J., Axmacher, N., Do Lam, A.T., Fell, J., Elger, C.E., Singer, W., & Melloni, L. Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. **J. Neurosci.**, v. 32, n. 43, p. 140909-14914. 2012.

Baars, B.J. **A cognitive theory of consciousness**. New York: Cambridge University Press, 1988.

_____. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. **Prog. Brain Res.**, v. 150, p. 45-53. 2005.

_____. **In the Theater of Consciousness: The Workspace of the Mind**. New York: Oxford University Press, 1997.

_____. Metaphors of consciousness and attention in the brain. **Trends Neurosci.**, v. 21, n. 2, p. 58-62. 1998.

_____. The conscious access hypothesis: origins and recent evidence. **Trends Cogn. Sci.**, v. 6, n. 1, p. 47-52. 2002.

Baars, B.J., & Franklin, S. How conscious experience and working memory interact. **Trends Cogn. Sci.**, v. 7, n. 4, p. 166-172. 2003.

Baddeley, A. The episodic buffer: a new component of working memory. **Trends Cogn. Sci.**, v. 4, n. 11, p. 417-423. 2000.

Bandler, R., & Shipley, M.T. Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? **Trends Neurosci.**, v. 17, n. 9, p. 379-389. 1994.

Barrett, L.F. Are emotions natural kinds? **Perspect. Psychol. Sci.**, v. 1, n. 1, p. 28-58. 2006b.

_____. Solving the Emotion Paradox: Categorization and the Experience of Emotion. **Personality And Social Psychology Review**, v. 10, n. 1, p.20-46. 2006a.

_____. The theory of constructed emotion: an active inference account of interoception and categorization. **Social Cognitive And Affective Neuroscience**, v. 12, n. 1, p.1-23. 2017.

Barrett, L.F., & Russell, J.A. Independence and bipolarity in the structure of current affect. **Journal of Personality and Social Psychology**, v. 74, n. 4, p.967-984. 1998.

Barrett, L.F., & Satpute, A.B. Historical pitfalls and new directions in the neuroscience of emotion. **Neuroscience Letters**. In press. 2017.

Barrett, L.F., & Simmons, W.K. Intoreceptive predictions in the brain. **Nat. Rev. Neurosci.**, v. 16, n. 7, p. 419-429. 2015.

Barrett, L.F., Lindquist, K.A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., & Brennan, L.. Of mice and men: natural kinds of emotions in the mammalian brain? A Response to Panksepp and Izard. **Perspect. Psychol. Sci.**, v. 2, n. 3, p. 297-311. 2007.

Bechara, A., Damasio, A.R., Damasio, H., & Anderson, S.W. Insensitivity to future 46 consequences following damage to human prefrontal cortex. **Cognition**, v. 50, n. 1-3, p. 7-15. 1994.

Bechara, A., Damasio, H., Damasio, A.R., & Lee, G.P. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. **J. Neurosci.**, v. 19, n. 13, p. 5473-5481. 1999.

Bechara, A., Tranel, D., & Damasio, H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. **Brain**, v. 123, p. 2189-2202. 2000.

Bickle, J., Mandik, P., & Landreth, A. The Philosophy of Neuroscience. **The Stanford Encyclopedia of Philosophy** (Summer 2012 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2012/entries/neuroscience/>>.

Block, N. Consciousness, accessibility, and the mesh between psychology and neuroscience. **Behavioral and Brain Sciences**, v. 30, n. 5-6, p.481-548. 2007.

Bornemann, B., Winkielman, P., & van der Meer, E. Can you feel what you do not see? Using internal feedback to detect briefly presented emotional stimuli. **Int. J. Psychophysiol.**, v. 85, p. 116-124. 2012.

Bosse, T., Jonker, C.M., & Treur, J. Formalization of Damasio's theory of emotion, feeling and core consciousness. **Conscious. Cogn.**, v. 17, n. 1, p. 94-113. 2008.

Bushnell, M.C., Goldberg, M.E., & Robinson, D.L. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. **J. Neurophysiol.**, v. 46, n. 4, p. 755-772. 1981.

Celesia, G.G. Visual perception and awareness: a modular system. **J. Psychophysiol.**, v. 24, p. 62-67. 2010.

Churchland, P.M. **Matter and Consciousness**. 3^a ed. Cambridge: The MIT Press, 2013.

Churchland, P.S. A neurophilosophical slant on consciousness research. **Prog. Brain Res.**, v. 149, p. 285-293. 2005.

_____. **Brain-Wise: studies in neurophilosophy**. Cambridge: The MIT Press. 2002.

_____. **Neurophilosophy: Toward a Unified Science of the Mind-Brain.** Cambridge: The MIT Press. 1986.

Craig, A.D. How do you feel? Interoception: the sense of the physiological condition of the body. **Nat. Rev. Neurosci.**, v. 3, n. 8, p. 655-666. 2002.

_____. How do you feel – now? The anterior insula and human awareness. **Nat. Rev. Neurosci.**, v. 10, n. 1, p. 59-70. 2009.

_____. Interoception: the sense of the physiological condition of the body. **Curr. Opin. Neurobiol.**, v. 13, n. 4, p. 500-505. 2003.

Crick, F., & Koch, C. A framework for consciousness. 1 **Nat. Neurosci.**, v. 6, n. 2, p. 119-126. 2003.

_____. Are we aware of neural activity in the primary visual cortex? **Nature**, v. 375, n. 6527, p. 121-123. 1995.

_____. Consciousness and neuroscience. **Cerebral Cortex**, v. 8, n. 2, p. 97-107. 1998.

_____. Towards a neurobiological theory of consciousness. **Seminars in The Neuroscience**, v. 2, p. 263-275. 1990.

_____. What is the function of the claustrum? **Philos. Trans. R. Soc. Lond. B Biol. Sci.**, v. 360, n. 1458, p. 1271-1279. 2005.

Damasio, A.R. **Descartes' Error: Emotion, Reason and the Human Brain.** New York: Putnam, 1994.

_____. **Looking for Spinoza: Joy, Sorrow, and the Feeling Brain.** Londres: William Heinemann, 2003.

_____. **Self Comes to Mind: Constructing the Conscious Brain.** New York: Pantheon, 2010.

_____. **The Feeling of What Happens: Body and Emotion in the Making of Consciousness.** San Diego: Harcourt, 1999.

Damasio, A., & Carvalho, G.B. The nature of feelings: evolutionary and neurobiological origins. **Nature Reviews Neuroscience**, v. 14, n. 2, p.143-152. 2013.

Damasio, A., Damasio, H., & Tranel, D. Persistence of feelings and sentience after bilateral damage of the insula. **Cerebral Cortex**, v. 23, n. 4, p. 833-846. 2013.

Damasio, A.R., Everitt, B.J., & Bishop, D. The somatic marker hypothesis and the possible functions of the prefrontal cortex [and discussion]. **Philos. Trans. R. Soc. Lond. B Biol. Sci.**, v. 351, n. 1346, p; 1413-1420. 1996.

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J., & Hichwa, R.D. Subcortical and cortical brain activity during the feeling of self-generated 37 emotions. **Nat. Neurosci.**, v. 3, n. 10, p. 1049-1056. 2000.

Dehaene, S. **Consciousness and the brain**: Deciphering how the brain codes our thoughts. New York: Viking Penguin, 2014.

Dehaene, S., & Changeux, J.P. Experimental and theoretical approaches to conscious processing. **Neuron**, v. 70, n. 2, p. 200-227. 2011.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., ... Le Bihan, D. Imaging unconscious semantic priming. **Nature**, v. 395, n. 6702, p. 597-600. 1998.

Dehaene, S., & Naccache, L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. **Cognition**, v. 79, n. 1, p. 1-37. 2001.

Dennett, D. Are we explaining consciousness yet? **Cognition**, v. 79, p. 221-237, 2001.

_____. **Consciousness Explained**. Boston: Little, Brown and Co., 1991.

Dijksterhuis, A., Bos, M.W., Nordgren, L.F., & van Baaren, R.B. On Making the Right Choice: The Deliberation-Without-Attention Effect. **Science**, v. 311, n. 5763, p.1005-1007. 2006.

Dubin, A. E., & Patapoutian, A. Nociceptors: the sensors of the pain pathway. **Journal Of Clinical Investigation**, v. 120, n. 11, p.3760-3772. 2010.

Duncan, S., & Barrett, L.F. Affect is a form of cognition: A neurobiological analysis. **Cognition & Emotion**, v. 21, n. 6, p.1184-1211. 2007.

Ekman, P. An argument for basic emotions. **Cogn. Emot.**, v. 6, n. 3–4, p. 169-200. 1992.

Ekman, P., Levenson, R.W., & Friesen, W.V. Autonomic nervous system activity distinguishes among emotions. **Science**, v. 221, n. 4616, p. 1208-1210. 1983.

Elston, G.N. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. **Cereb. Cortex**, v. 13, p. 1124–1138. 2003.

Ester, E.F., Rademaker, R.L., & Sprague, T.C. How do visual and parietal cortex contribute to visual short-term memory? **eNeuro**, v. 3, n. 2, ENEURO.0041-16.2016. 2016b.

Ester, E.F., Sutterer, D.W., Serences, J.T., & Awh, E. Feature-selective attentional modulations in human frontoparietal cortex. **J. Neurosci.**, v. 36, n. 31, p. 8188-8199. 2016a.

Etkin, A., & Wager, T. D. Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia. **American Journal Of Psychiatry**, v. 164, n. 10, p.1476-1488. 2007.

Feinstein, J.S., Buzza, C., Hurlemann, R., Follmer, R.L., Dahdaleh, N.S., Coryell, W.H., ... Wemmie, J.A. Fear and panic in humans with bilateral amygdala damage. **Nat. Neurosci.**, v. 16, p. 270–272. 2013.

Feinstein, J.S., Khalsa, S.S., Salomons, T.V., Prkachin, K.M., Frey-Law, L.A., Lee, J.E., ... Rudrauf, D. Preserved emotional awareness of pain in a patient with extensive bilateral damage to the insula, anterior cingulate, and amygdala. **Brain Struct. Funct.**, v. 221, n. 3, 1499-1511. 2016.

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. Binocular rivalry: frontal activity relates to introspection and action but not to perception. **J. Neurosci.**, v. 34, n. 5, p. 1738-1747. 2014.

Fresco, N. The Explanatory Role of Computation in Cognitive Science. **Minds And Machines**, v. 22, n. 4, p.353-380, 13. 2012.

Frijda, N.H. **The Emotions**. Cambridge: Cambridge University Press, 1986.

Friston, K.J. et al. Event-Related fMRI: Characterizing Differential Responses. **Neuroimage**, v. 7, n. 1, p.30-40, jan. 1998.

Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., ... Naccache, L. Converging intracranial markers of conscious access. **PLoS Biology**, v. 17, n. 3, e61. 2009.

Grafton, S.T. et al. Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. **Journal of Neuroscience**, v. 12, n. 7, p. 2542-2548. 1992.

Gray, C.M., König, P., Engel, A.K., & Singer, W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. **Nature**, v. 338, n. 6213, p. 334-337. 1989.

Grill-Spector, K., & Malach, R. The human visual cortex. **Annu. Rev. Neurosci.**, v. 27, p. 649–677. 2004.

Harmon-Jones, E., Gable, P.A., & Price, T.F. Narrow and positive affect always broaden the mind? Considering the influence of motivational intensity on cognitive scope. **Curr. Dir. Psychol. Sci.**, v. 22, p. 301–307. 2013.

Jackson, F. Epiphenomenal Qualia. **Philosophical Quarterly**, v. 32, n. 127, p.127-136. 1982.

James, W. What is an emotion? **Mind**, v. 9, n. 34, p. 188-205. 1884.

Keltner, D., & Gross, J.J. Functional Accounts of Emotions. **Cognition & Emotion**, v. 13, n. 5, p.467-480. 1999.

Koch, C., Massimini, M., Boly, M., & Tononi, G. Neural correlates of consciousness: progress and problems. **Nat. Rev. Neuroscience**, v. 17, n. 5, p. 307-321. 2016.

Kreibig, S.D. Autonomic nervous system activity in emotion: A review. **Biological Psychology**, v. 84, n. 3, p.394-421. 2010

Koch, C., & Tsuchiya, N. Attention and consciousness: two distinct brain processes. **Trends Cogn. Sci.**, v. 11, n. 1, p. 16-22. 2007.

Ledoux, J.E. **Anxious: Using the Brain to Understand and Treat Fear and Anxiety**. New York: Viking, 2015.

_____. Emotion, Memory and the Brain. **Scientific American**, v. 270, n. 6, p. 50-57. 1994.

_____. Rethinking the emotional brain. **Neuron**, v. 73, n. 4, p. 653-676. 2012.

_____. **The Emotional Brain**. New York: Simon and Schuster, 1996.

LeDoux, J.E., & Brown, R. A higher-order theory of emotional consciousness. **Proc. Natl. Acad. Sci. U.S.A.**, v. 114, p. E2016–E2025. 2017.

LeDoux, J.E., & Pine, D.S. Using neuroscience to help Understand fear and anxiety: a two-system framework. **Am. J. Psychiatry**, v. 173, p. 1083–1093. 2016.

Levenson, R.W.. The Autonomic Nervous System and Emotion. **Emotion Review**, v. 6, n. 2, p.100-112. 2014.

Libet, B., Gleason, C.A., Wright, E.W., & Pearl, D.K. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. **Brain**, v. 106, p. 623–642. 1983.

Llinás, R., & Ribary, U. Coherent 40-Hz oscillation characterizes dream state in humans. **Proc. Natl. Acad. Sci. U. S. A.**, v. 90, n. 5, p. 2078-2081. 1993.

London, B.M. et al. Electrical Stimulation of the Proprioceptive Cortex (Area 3a) Used to Instruct a Behaving Monkey. **IEEE Transactions On Neural Systems And Rehabilitation Engineering**, v. 16, n. 1, p.32-36. 2008.

Mackintosh, N.J. A theory of attention: variations in the associability of stimuli with reinforcement. **Psychological Review**, v. 82, n. 4, p. 276-298. 1975.

Marcel, A.J. Conscious and unconscious perception: experiments on visual masking and word recognition. **Cogn. Psychol.**, v. 15, n. 2, p. 197-237. 1983.

Magnin, M., Rey, M., Bastuji, H., Guillemant, P., Mauguière, F., & Garcia-Larrea, L. Thalamic deactivation at sleep onset precedes that of the cerebral cortex in humans. **Proc. Natl. Acad. Sci. U. S. A.**, v. 107, n. 8, p. 3829-3833. 2010.

Meyer, K. Primary sensory cortices, top-down projections and conscious experience. **Prog. Neurobiol.**, v. 94, n. 4, p. 408-417. 2011.

Miller, E.K., & Cohen, J.D. An Integrative Theory of Prefrontal Cortex Function. **Annual Review of Neuroscience**, v. 24, n. 1, p.167-202. 2001.

Milner, P.M. A model for visual shape recognition. **Psychological Review**, v. 81, n. 6, p. 521-535. 1974.

Morris, J.S., DeGelder, B., Weiskrantz, L., & Dolan, R.J. Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. **Brain**, v. 124, p. 1241–1252. 2001.

Naccache, L., & Dehaene, S. Reportability and illusions of phenomenality in the light of the global neuronal workspace model. **Behavioral and Brain Sciences**, v. 30, n. 5-6, p. 518-520. 2007.

Neisser U. **Cognitive psychology**. New York: Appleton-Century-Crofts, 1967.

Noback, C.R. et al. **The human nervous system: structure and function**. Totowa: Human Press Inc., 2005.

Panksepp, J. Affective consciousness: Core emotional feelings in animals and humans. **Consciousness And Cognition**, v. 14, n. 1, p.30-80. 2005.

_____. Affective consciousness in animals: perspectives on dimensional and primary process emotion approaches. **Proceedings Of The Royal Society B: Biological Sciences**, v. 277, n. 1696, p.2905-2907. 2010.

_____. Affective neuroscience of the emotional BrainMind: evolutionary perspectives and implications for understanding depression. **Dialogues in Clinical Neuroscience**, v. 12, n. 4, p. 533-545. 2010b.

_____. **Affective Neuroscience: The Foundations of Human and Animal Emotions**. New York: Oxford University Press, 1998.

_____. Emotions as natural kinds within the mammalian brain. In M. Lewis e J.M. Haviland-Jones (Eds.), **Handbook of emotions**, 2^a ed., p. 137–156. New York: Guilford, 2000.

_____. Neurologizing the psychology of affects: How appraisal-based constructivism and basic emotion theory can coexist. **Perspect. Psychol. Sci.** v. 2, n. 3, p. 281-296. 2007.

Panksepp, J., Lane, R.D., Solms, M., & Smith, R. Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. **Neuroscience & Biobehavioral Reviews**, v. 76, p.187-215. 2017.

Parvizi, J., Van Hoesen, G.W., Buckwalter, J., & Damasio, A. Neural connections of the posteromedial cortex in the macaque. **Proc. Natl. Acad. Sci. U. S. A.**, v. 103, n. 5, p. 1563-1568. 2006.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L.B. The plastic human brain cortex. **Annu. Rev. Neurosci.**, v. 28, p. 377–401. 2005.

Paul, E.S., Harding, E.J., & Mendl, M. Measuring emotional processes in animals: the utility of a cognitive approach. **Neuroscience & Biobehavioral Reviews**, v. 29, n. 3, p.469-491. 2005.

Penfield, W., & Boldrey, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. **Brain**, v. 60, p. 389–443. 1937.

Place, U. T. Is consciousness a brain process? **The Mind-brain Identity Theory**, p.42-51. 1970.

Pomerantz, J.R. Perception: overview. In L. Nadel (Ed.). **Encyclopedia of cognitive science** (p. 527–537). London: Nature Publishing Group, 2003.

Posner, J. et al. The neurophysiological bases of emotion: An fMRI study of the affective circumplex using emotion-denoting words. **Human Brain Mapping**, v. 30, n. 3, p.883-895. 2009.

Price, D. D. Characteristics of second pain and flexion reflexes indicative of prolonged central summation. **Experimental Neurology**, v. 37, n. 2, p.371-387. 1972.

Robbins, T.W., & Everitt, B.J. Neurobehavioural mechanisms of reward and motivation. **Current Opinion in Neurobiology**, v. 6, n. 2, p.228-236. 1996.

Russell, J.A.. Core affect and the psychological construction of emotion. **Psychological Review**, v. 110, n. 1, p.145-172. 2003.

Saalman, Y.B., & Kastner, S. The cognitive thalamus. *Frontiers In Systems Neuroscience*, v. 9, artigo 39. 2015.

Safavi, S., Kapoor, V., Logothetis, N.K., & Panagiotaropoulos, T.I. Is the frontal lobe involved in conscious perception? **Front. Psychol.**, v. 5, artigo 1063. 2014.

Shannon, C.E., & Weaver, W. **The Mathematical Theory of Communication**. Urbana: The University of Illinois Press, 1949.

Smart, J. J. C. Sensations and Brain Processes. **The Philosophical Review**, v. 68, n. 2, p.141-156. 1959.

Sperling, G. The information available in brief visual representation. **Psychol. Monogr.**, v. 74, n. 11, p. 1-29. 1960.

Tamietto, M., & de Gelder, B. Neural bases of the non-conscious perception of emotional signals. **Nat. Rev. Neurosci.**, v. 11, p. 697–709. 2010.

Teixeira, J.F. **Mente, Cérebro e Cognição**. 4ª ed. Petrópolis: Vozes, 2000.

Thagard, P. Cognitive Science. **The Stanford Encyclopedia of Philosophy** (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2014/entries/cognitive-science/>>. 2012.

Tononi, G. An information integration theory of consciousness. **BMC Neurosci.**, v. 5, artigo 42. 2004.

_____. Consciousness as integrated information: a provisional manifesto. **Biol. Bull.**, v. 215, n. 3, p. 216-242. 2008.

Tononi, G., Boly, M., Massimini, M., & Koch, C. Integrated information theory: from consciousness to its physical substrate. **Nat. Rev. Neurosci.**, v. 17, n. 7, p. 450-461. 2016.

Tononi, G., & Koch, C. Consciousness: here, there and everywhere? **Philos. Trans. R. Soc. Lond. B Biol. Sci.**, v. 370, n. 1668, artigo 20140167. 2015.

Tye, M. **Ten Problems of Consciousness**. Cambridge: The MIT Press. 1995.

Velmans, M. **Understanding Consciousness**. 2ª edição. London: Routledge, 2009.

Whittingstall, K., & Logothetis, N.K. Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. **Neuron**, v. 64, n. 2, p. 281-289. 2009.

Williams, L.M., Das, P., Liddell, B.J., Kemp, A.H., Rennie, C.J., & Gordon, E. Mode of functional connectivity in amygdala pathways dissociates level of awareness for signals of fear. **J. Neurosci.**, v. 26, p. 9264–9271. 2006.

Wilson-mendenhall, C.D., Barrett, L.F., & Barsalón, L.W. Neural Evidence That Human Emotions Share Core Affective Properties. **Psychological Science**, v. 24, n. 6, p.947-956. 2013.

Wojciulik, E., & Kanwisher, N. The generality of parietal involvement in visual attention. **Neuron**, v. 23, n. 4, p. 747-764. 1999.

Zajonc, R. B. On the primacy of affect. **Am. Psychol.**, v. 39, p. 117–123. 1984.