

**UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE FILOSOFIA E CIÊNCIAS SOCIAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA**

Como Atribuir Consciência aos Animais

Victor Machado Barcellos

Rio de Janeiro

2022

VICTOR MACHADO BARCELLOS

Como Atribuir Consciência aos Animais

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio de Janeiro, como requisito para a obtenção do título de Mestre em Filosofia.

Orientador: Prof. Dr. Rodrigo A. S. Gouvea
Coorientador: Prof. Dr. Roberto H. S. Pereira

Rio de Janeiro

2022

CIP - Catalogação na Publicação

B242c Barcellos, Victor Machado
 Como Atribuir Consciência aos Animais / Victor
Machado Barcellos. -- Rio de Janeiro, 2022.
 186 f.

 Orientador: Rodrigo Azevedo dos Santos Gouvea.
 Coorientador: Roberto Horácio de Sá Pereira.
 Dissertação (mestrado) - Universidade Federal do
Rio de Janeiro, Instituto de Filosofia e Ciências
Sociais, Programa de Pós-Graduação em Filosofia,
2022.

 1. Consciência. 2. Consciência Animal. 3.
Filosofia da Consciência. 4. Ciência da
Consciência. 5. Cognição Animal. I. Gouvea, Rodrigo
Azevedo dos Santos, orient. II. Pereira, Roberto
Horácio de Sá, coorient. III. Título.

Victor Machado Barcellos

COMO ATRIBUIR CONSCIÊNCIA AOS ANIMAIS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio de Janeiro, como parte dos requisitos para obtenção do título de Mestre em Filosofia.

Áreas de Concentração: Filosofia da Mente, Filosofia da Consciência e Filosofia da Ciência da Consciência.

Aprovada por:

Presidente, Prof. Dr. Rodrigo Azevedo dos Santos Gouvea (orientador) – UFRJ

Prof. Dr. Roberto Horácio de Sá Pereira (coorientador) – UFRJ

Prof. Dr. Marco Aurélio Sousa Alves – UFSJ

Prof. Dr. Sérgio Farias de Souza Filho – UFRPE

Rio de Janeiro, 02 de setembro de 2022.

Às minhas avós (*in memoriam*)

AGRADECIMENTOS

Agradeço aos meus orientadores, profs. Rodrigo Gouvea e Roberto Horácio, pela paciência e atenção que tiveram durante todo o processo de pesquisa. Sou imensamente grato pelas inúmeras discussões que tivemos e por todo o aprendizado que pude retirar dessas conversas. Sem dúvida alguma as críticas e sugestões de ambos foram vitais para o desenvolvimento deste trabalho.

Agradeço também ao prof. Sérgio Farias pela sua imprescindível colaboração durante o processo de pesquisa. Devo dizer que muito da inspiração naturalista deste trabalho vem de sua positiva influência.

Expresso também a minha gratidão ao prof. Marco Aurélio por ter feito críticas e sugestões importantes a versões anteriores de alguns dos capítulos. Espero ter em alguma medida melhorado esses pontos.

Agradeço aos profs. Peter Carruthers e Peter Godfrey-Smith por gentilmente terem esclarecido algumas de minhas dúvidas.

Sou grato também à minha família e à minha companheira Beatriz Aguiar por ter me apoiado imensamente durante este árduo período de pesquisa. Obrigado Bea por ter lido uma parte da dissertação e ter tecido diversos apontamentos pertinentes.

Diversos amigos e colegas me apoiaram e colaboram de alguma forma durante a pesquisa. Dentre eles cito: Allan Bruno, Gustavo de Souza, Maria Amélia, Marllon Costa, Nicholas Ferreira, Raphael Torres, Ricardo Villar, Rodolpho Nagel e Victor Gomes. Obrigado a todos!

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de pesquisa que possibilitou a realização deste trabalho durante a pandemia de COVID-19.

RESUMO

Os debates sobre a existência da consciência em animais não humanos vêm ganhando cada vez mais destaque nos círculos científicos e filosóficos ao redor do mundo. Um importante empecilho que a área atualmente enfrenta diz respeito à construção de um método confiável capaz de identificar se um determinado animal não humano possui estados mentais conscientes. A literatura nomeia essa questão de “o problema da mensuração da consciência animal”. A presente dissertação possui como objetivo analisar e responder esse problema através da combinação entre o uso de evidências empíricas e de ferramentas filosóficas. Dito isso, existem inúmeras formas de respondê-lo. Ao longo do texto, analisarei algumas das que são consideradas mais relevantes na discussão atual. Uma delas é a concepção cética de que não há questão de fato sobre os animais não humanos serem ou não conscientes. Em contraste, as abordagens neutras e newtonianas buscam se esquivar de compromissos metafísicos e teóricos para atribuir consciência aos animais. Ao final da dissertação defenderei a chamada “abordagem teórica leve”, posição originalmente proposta pelo filósofo Jonathan Birch. Buscarei mostrar que ela é atualmente a hipótese empiricamente mais adequada e filosoficamente plausível dentre as disponíveis na literatura.

Palavras-Chave: Consciência; Consciência Animal; Ciência da Consciência; Filosofia da Consciência; Cognição Animal.

ABSTRACT

Debates concerning the existence of consciousness in non-human animals are gaining more and more prominence in scientific and philosophical circles around the world. An important obstacle that the area currently faces concerns the construction of a reliable method capable of identifying whether a certain non-human animal has conscious mental states. The literature names this issue “the measurement problem of animal consciousness”. The present dissertation aims to analyze and answer this problem through the combination of the empirical evidence and philosophical tools. That said, there are numerous ways to answer it. Throughout the text, I will analyze some of those that are considered most relevant in the current discussion. One is the skeptical view that there is no fact of the matter regarding the presence of consciousness in non-human animals. In contrast, theory neutral and Newtonian approaches seek to evade metaphysical and theoretical commitments to ascribing consciousness to non-human animals. At the end of the dissertation, I will defend the so-called “theory-light approach”, a position originally proposed by the philosopher Jonathan Birch. I will try to show that it is currently the most empirically adequate and philosophically plausible hypothesis among those available in the literature.

Keywords: Consciousness; Animal Consciousness; Science of Consciousness; Philosophy of Consciousness; Animal Cognition.

“He who understands baboon would do more towards metaphysics than Locke.”

(Darwin, Charles)

LISTA DE FIGURAS

Figura 1 – Projeto de pesquisa interdisciplinar sobre a consciência animal	13
Figura 2 - Interpretação de Block sobre o experimento de Sperling.....	28
Figura 3 – O Córtex pré-frontal.....	30
Figura 4 – As principais regiões do cérebro humano.....	32
Figura 5 – A semelhança estatística	34
Figura 6 – Interpretação da abordagem fragmentada da fenomenologia sobre o experimento de Sperling	35
Figura 7 - O primeiro questionário de Cova et al.....	36
Figura 8 - O segundo questionário de Cova et al.....	36
Figura 9 – Representação do espaço de trabalho global.....	63
Figura 10 – Representação da GWT funcionando no cérebro humano.....	64
Figura 11 – Comparação entre os estados cerebrais durante a percepção consciente e inconsciente de um dado estímulo.....	67
Figura 12 – Experimento sobre a consciência perceptual em corvos	111
Figura 13 – Imagem de fMRI da cabeça de uma criança hidranencéfala.....	128
Figura 14 – Representação aproximada do campo visual cego de um indivíduo portador de visão cega.....	131
Figura 15 – Fluxo ventral e dorsal.	135
Figura 16 – Condicionamento de rastreo	143
Figura 17 - Condicionamento com <i>delay</i> encurtado.....	143
Figura 18 – Modelo do experimento de Ben-Haim et al utilizando um estímulo incongruente.....	145
Figura 19 - O experimento de Ben-Haim et al utilizando um estímulo incongruente em macacos.....	146
Figura 20 – Árvore filogenética.....	154
Figura 21 – Evolução dos olhos	155
Figura 22 – Representação gráfica do modelo gradualista <i>scala naturae</i> da consciência	160
Figura 23 – Representação gráfica do modelo gradualista multidimensional da consciência.....	160

SUMÁRIO

INTRODUÇÃO	13
1 O QUE É A CONSCIÊNCIA?	19
1.1 Explicitando a Consciência	19
1.2 O Debate do Transbordamento (<i>The Overflow Debate</i>)	24
1.2.1 Interpretações Contrárias.....	32
1.3 Os Problemas da Distribuição e da Mensuração da Consciência Animal.....	38
1.3.1 A Abordagem da Percepção Direta	41
1.3.2 A Abordagem Teórica	48
1.3.3 A Abordagem Epistêmica	51
2 DEVEMOS NOS IMPORTAR COM A CONSCIÊNCIA ANIMAL!	54
2.1 Introdução	54
2.2 A Consciência como um Fenômeno Tudo-ou-Nada.....	56
2.2.1 Realismo de <i>qualia</i> e Irrealismo de <i>qualia</i>	58
2.3 A Teoria do Espaço de Trabalho Global	61
2.4 A Estratégia do Conceito Fenomenal.....	68
2.5 A Motivação por trás da Postura Cética de Carruthers	76
2.6 O Argumento Semântico Negativo	82
2.7 O Argumento Semântico Positivo	84
2.8 Objeções a Carruthers.....	88
2.8.1 O Argumento do Cientista Alienígena.....	89
2.8.2 A Semântica Intencional Não Naturalizada.....	93
2.8.3 Seria a Consciência um Fenômeno Tudo-ou-Nada?	97
3 A ESTRATÉGIA DOS MARCADORES DE CONSCIÊNCIA	104
3.1 Introdução	104
3.2 Distinções e Clarificações	105

3.2.1	Inferência para a Melhor Explicação	105
3.2.2	Correlatos Neurais, Marcadores e Constituintes da Consciência.....	109
3.3	A Estratégia dos Marcadores de Consciência	112
3.4	A Abordagem Padrão	115
3.5	A Abordagem Newtoniana de Tye	120
3.5.1	O Argumento da Inevitabilidade da Teoria	127
3.5.2	O Desafio de Carruthers-Dawkins.....	132
3.6	A Abordagem Teórica Leve	138
3.6.1	Respostas a Possíveis Objeções.....	147
CONCLUSÃO.....		158
APÊNDICE		160
REFERÊNCIAS		161

INTRODUÇÃO

Presenciamos atualmente um verdadeiro *boom* de pesquisas sobre consciência em animais não humanos¹. Anteriormente o tópico causava um certo incômodo entre os cientistas e, em menor grau, entre os filósofos. Isto pode ser verificado através do tratamento marginal ou mesmo excludente que a consciência recebia nos livros-texto sobre comportamento e cognição animal (ALCOCK e RUBENSTEIN, 2019; KAPPELER, 2021; SHETTLEWORTH, 2010; WYNNE e UDELL, 2020). No entanto, uma quantidade cada vez maior de filósofos e cientistas de diferentes campos vem se engajando no tema, gerando conhecimento e produzindo debates que são imprescindíveis para o desenvolvimento dessa nova área interdisciplinar denominada de “Pesquisas sobre Consciência Animal” (Figura 1) (BIRCH et al, 2022). Cito alguns dentre os principais nomes da área: a dupla de biólogas Simona Ginsburg e Eva Jablonka (2019), a dupla de neurocientistas Todd E. Feinberg e Jon M. Mallatt (2017, 2018), bem como os filósofos Jonathan Birch (2020a, 2022a), Michael Tye (2017), Peter Carruthers (2019) e Peter Godfrey-Smith (2019a, 2020a).



Figura 1 – Projeto de pesquisa interdisciplinar sobre a consciência animal, incluindo áreas como: ciência do bem-estar animal, biologia evolutiva, filosofia, psicologia comparada e neurociência cognitiva. Birch et al, 2022, p. 10.

¹ Doravante, empregarei o termo “animal”, salvo quando for indicado.

Somado a isso, existem outros sinais que indicam o interesse crescente em relação ao tema. Pode-se citar a criação da revista interdisciplinar *Animal Sentience* e do acordo denominado “A Declaração de Cambridge sobre a Consciência”, assinado por diversos pesquisadores, dentre eles o célebre físico britânico Stephen Hawking (LOW et al, 2012). Ademais, uma importante consequência do destaque que o tema vem ganhando no debate público é o recente reconhecimento de *seres sencientes* que diversas espécies de invertebrados, mamíferos e aves vem recebendo por parte das legislações de diversos países ao redor do mundo (BROWNING e VEIT, 2022; ROWAN et al, 2022). No Reino Unido, por exemplo, os cefalópodes e os crustáceos decápodes foram recentemente reconhecidos como seres sencientes na legislação sobre o assunto (REINO UNIDO, 2022). É importante salientar que o empreendimento de convencimento do governo britânico foi feito a partir de trabalhos robustos realizados por cientistas e filósofos empiricamente bem-informados (BIRCH et al, 2021; CRUMP et al, 2022; de WAAL e ANDREWS, 2022).

Do tópico da consciência animal emergem diversas questões importantes. Nesta dissertação me restringirei a dois problemas que considero estarem intimamente interligados. O primeiro é o problema da distribuição da consciência animal que diz respeito à extensão da consciência na árvore filogenética entre as diferentes linhagens de animais (BAYNE, 2021, p. 208). O segundo é o problema da mensuração da consciência animal que aborda qual seria o método adequado para indicar a presença da consciência nos animais (BROWNING e VEIT, 2020). O motivo pelo qual os dois problemas estão conectados é que para identificarmos a extensão da consciência no reino animal devemos possuir um método capaz de indicar se uma determinada população de animais possui ou não estados mentais conscientes.

Acredito ser de suma importância científica, filosófica e ética oferecer respostas aos problemas da distribuição e da mensuração. Digo científico e filosófico na medida em que cada uma das inúmeras teorias da consciência disponíveis, sejam elas de natureza neurobiológica ou especulativa, implicam em um rol diferente de animais considerados conscientes (SETH e BAYNE, 2022). Algumas teorias são mais inclusivas, abarcando não somente os hominínios², mas também linhagens filogeneticamente distantes da nossa, como as dos artrópodes e outros invertebrados (BARRON e KLEIN, 2016; GODFREY-SMITH, 2020a; KOCH, 2020; TYE, 2021). Ao passo que outras teorias restringem de forma contundente as espécies consideradas conscientes (CARRUTHERS, 2000; GRAZIANO, 2019; MACPHAIL, 2008; PRINZ, 2018).

² Empregarei o termo “hominínio” para se referir aos membros da linhagem humana e “hominídeo” aos grandes primatas, humanos e seus ancestrais.

Assim, o fornecimento de respostas a esses dois problemas ajudaria a determinar qual é a teoria da consciência mais provável de ser a verdadeira. Ora, se o conjunto de animais dotado de consciência for amplo, abarcando organismos como os artrópodes e cefalópodes, isto constituiria uma forte evidência a favor das teorias inclusivas da consciência. Por outro lado, se o conjunto de animais conscientes for restrito a um pequeno grupo, como apenas o dos grandes primatas, isto constituiria uma evidência a favor das teorias restritivas da consciência. Como Dung (2022, p. 3) bem expressa, a resolução desses problemas pode servir como um teste para saber qual dentre as várias teorias da consciência existentes é a correta. Ademais, a solução desses problemas também ajudaria a resolver os casos em disputa de humanos, como os de pacientes em estado vegetativo e comatoso, e dos indivíduos com condições congênitas severas, como as das crianças portadoras de hidranencefalia.

Do ponto de vista da ética, a importância se dá na medida em que diversas teses éticas sustentam que a consciência, ou ao menos a experiência consciente dotada de valência (KOSGAARD, 2018; SINGER, 2010), é uma condição necessária, suficiente e/ou moralmente relevante para o estabelecimento do status moral de um certo organismo (BIRCH, 2022a). Em outras palavras, para que um determinado animal seja considerado digno de relevância moral, segundo essas teses, ele deverá ser portador de estados mentais conscientes.³ Assim, ao responder os problemas da distribuição e da mensuração conseguiríamos saber quais são os animais conscientes e, por consequência, os animais que são moralmente relevantes. A ciência do bem-estar animal seria fortemente impactada, gerando efeitos imediatos nas tomadas de decisões das políticas públicas sobre o bem-estar desses seres (BROWNING, 2022; YEATES, 2022). Práticas que desnecessariamente causam sofrimento, como o cozimento de lagostas vivas e a aquicultura de polvos, poderiam ser consideradas moralmente erradas e ilegais. Isto posto, esquematizei a dissertação da seguinte forma.

O 1º capítulo consiste em uma preparação de terreno. Em primeiro lugar, buscarei estabelecer o que entendi por “consciência”. Como se encontra razoavelmente bem

³ Alguns autores, mais notavelmente a bióloga Marian Dawkins (2017, 2021), defendem que não se deve definir o bem-estar animal através dos termos “consciência” e “senciência” (cf. também CARRUTHERS, 2019). Dawkins (2017, 2021) expõe uma série de argumentos para sustentar tal posição, como: a suposta carência de evidências e entendimento que a ciência atual possui sobre quais são os mecanismos cerebrais que produzem e sustentam estados mentais conscientes e como eles o fazem, a falta de um consenso científico sobre qual é a teoria da consciência correta, a ausência de métodos confiáveis que distingam estados afetivos conscientes dos inconscientes, dado que ambos são capazes de gerar comportamentos similares, entre outros. Segundo Dawkins (2022), estabelecer políticas públicas sobre o bem-estar animal apelando à consciência/senciência seria não só um erro, mas um desserviço às pesquisas científicas rigorosas sobre a consciência em animais humanos e não humanos. Estaríamos, segundo a autora, contaminando a ciência da consciência com vieses e ideologias políticas para que de forma apressada diversos animais fossem reconhecidos como sencientes. Quanto às respostas aos argumentos de Dawkins, cf. Birch, 2022b, no prelo; Crump, 2022.

estabelecido na literatura, “consciência” é um conceito ambíguo e difícil, se não impossível, de se definir objetivamente. Deste modo, explicitarei o termo “consciência” por meio do uso de exemplos. Creio que atualmente essa seja, ao lado da locução “há algo para alguém que é como ser x ”, onde “ x ” é uma variável que só pode ser substituída por estados mentais conscientes, a melhor forma disponível de explicitação do conceito “consciência” (NAGEL, 1974). Realizarei também distinções importantes entre os diversos significados que “consciência” possui na literatura, como: consciência de criatura transitiva e intransitiva, consciência fenomenal e consciência de acesso. Em seguida, apresentarei de forma panorâmica a controvérsia acerca da possível distinção empírica entre a consciência fenomenal e a consciência de acesso através do debate do transbordamento. Procurarei apresentar alguns dos argumentos a favor do transbordamento da consciência fenomenal sob a de acesso, assim como certos argumentos contrários. Posteriormente, explicarei do que se trata os problemas da distribuição e da mensuração para, por fim, apresentar as três principais tentativas de respondê-las. São elas: a abordagens teórica, a epistêmica e a da percepção direta.

No 2º capítulo, apresentarei criticamente a posição radical do filósofo e cientista cognitivo Peter Carruthers (2018a, 2018b, 2019, 2020a). Segundo Carruthers, não existiria questão de fato acerca da consciência em animais não humanos. O motivo, segundo ele, é que na medida em que a melhor e mais completa teoria reducionista da consciência humana diz respeito a apenas como a consciência se realiza em uma arquitetura cognitiva como a dos humanos, a sua transposição para fora deste domínio se torna semanticamente indeterminado. De acordo com essa posição, devemos abandonar o empreendimento natimorto que é o da investigação da consciência em animais e direcionar às pesquisas para apenas a natureza e o funcionamento da cognição entre eles.

No 3º capítulo, me aprofundarei na abordagem epistêmica. Grosso modo, essa é uma posição que procura se esquivar de grandes comprometimentos teóricos e/ou metafísicos. Os proponentes desse tipo de abordagem utilizam largamente a inferência para a melhor explicação em suas hipóteses. Deste modo, apresentarei brevemente no que consiste o tipo de inferência utilizada. Em seguida, como forma de clarificar o funcionamento da abordagem epistêmica, apresentarei a tese em que ela se assenta, a saber, a chamada “estratégia dos marcadores de consciência”. Farei também uma importante distinção entre os marcadores de consciência, os correlatos neurais da consciência e os constituintes da consciência. Prosseguindo, analisarei duas posições epistêmicas influentes no debate: a abordagem padrão e a abordagem newtoniana de Michael Tye (2017, 2018). Por fim, defenderei a abordagem

teórica leve, tese que considero ser atualmente a mais plausível dentre as disponíveis na literatura (BIRCH, 2020a).

Quando se pensa em consciência animal, as pessoas frequentemente tendem a direcionar os seus interesses à ética. Embora considere que o presente trabalho possua consequências para a área, devo alertar que não abordarei os problemas relacionados a ela. O campo da ética animal é extremamente vasto e possui uma enorme literatura especializada. Seria necessária uma nova dissertação para fornecer um tratamento adequado sobre as consequências morais das teses aqui apresentadas. Porém, caso o leitor possua interesse no tópico, recomendo a leitura do manual escrito por Bob Fischer (2021). Esse livro apresentará um bom panorama da disciplina e indicará alguns caminhos possíveis para dar prosseguimento aos estudos que, porventura, interessarem ao leitor.

Isto posto, gostaria de brevemente expor a metodologia que adotarei nesta dissertação. Aqui, tentei na medida do possível realizar um trabalho filosófico que fosse empiricamente bem-informado e fundamentado nas evidências científicas disponíveis. Com efeito, contraponho essa postura à tradicional filosofia da poltrona. Entendo aqui “filosofia da poltrona” como aquela que é feita sem a utilização de dados empíricos e guiada inteiramente pelo raciocínio puro. Se segue analogamente desta concepção que assim como um trabalho não precisa ser experimental para ser considerado científico, *mutatis mutandis*, um trabalho de filosofia que não realiza experimentos não seria considerado como da poltrona. Em outras palavras, um trabalho de filosofia que se utiliza de dados empíricos, mas que não realiza experimentos, *não* seria considerado como filosofia da poltrona. A contraposição aqui realizada não é, portanto, entre a filosofia da poltrona, em seu sentido tradicional, e a filosofia experimental (*x-phi*) (KNOBE e NICHOLS, 2017; WILLIAMSON, 2021), mas entre a filosofia empiricamente informada/fundamentada e a filosofia puramente especulativa.

Isso pode ser melhor observado através da analogia fornecida por Carruthers (2015) entre a física experimental e teórica. Enquanto a primeira realiza experimentos, a segunda, em geral, utiliza os dados obtidos pela primeira para construir as suas teorias e modelos explicativos. Desta forma, um físico teórico que durante toda a sua vida nunca esteve em um laboratório para realizar sequer um experimento não seria descredenciado como cientista, embora faça ciência da poltrona de seu escritório. Evidentemente, nem sempre é clara a distinção entre o físico teórico e o experimental. Pense, por exemplo, nos pesquisadores que

trabalham com física computacional.⁴ Contudo, isso não elimina os casos que claramente são teóricos e outros que são experimentais. A única consequência notável disso seria a de que conforme nos aproximamos de suas fronteiras, nos deparamos com uma zona cinzenta onde é difícil delimitar com clareza se um empreendimento é completamente teórico ou experimental.

De forma análoga à visão apresentada acima, alguém que faz filosofia a partir de dados empíricos, mas que não vai a campo produzir experimentos, não seria considerado um filósofo da poltrona. Ora, aqui o que importa não é se o cientista ou o filósofo está trabalhando ou não da poltrona, mas *como* ele está trabalhando da poltrona. Utilizo “poltrona” apenas por questões pragmáticas, embora acredite que este não seja o termo mais adequado para se referir à filosofia especulativa tradicional.

Em suma, assim como Carruthers (2006, pp. xiv-xv) expressa, creio que atualmente a melhor forma de se fazer filosofia, ao menos a que trata da mente, é por meio da utilização *não ingênua* dos dados fornecidos pelas ciências.⁵ Diferentes autores propuseram metodologias filosóficas com essa postura, sendo algumas mais radicais ou moderadas do que outras. Por exemplo, Peter Godfrey-Smith (2009) sugeriu a chamada “filosofia da natureza”⁶, Eric Schliesser (2019) a “filosofia sintética” e Patricia Churchland (2022) a “neurofilosofia”. A proposta de Schliesser foi inclusive esboçada a partir de um exame metodológico de dois livros, sendo um deles o “*Outras Mentes: O Polvo e a Origem da Consciência*” de Godfrey-Smith (2019a).⁷ Por fim, espero que esta dissertação seja de alguma forma útil não só aos filósofos, mas também aos cientistas e a qualquer um que se interesse pelo tema.

⁴ Recentemente pesquisadores brasileiros realizaram pela primeira vez na história uma simulação do comportamento de um buraco negro através do uso de Inteligência Artificial, especificamente pela técnica de *Deep Learning* (DUARTE et al, 2022). Seria esse um trabalho de astrofísica teórica, experimental ou ambas? Creio ser difícil apontar com certeza em qual dessas opções o trabalho se encaixa.

⁵ Embora seja polêmico, acredito que, *mutatis mutandis*, essa postura possa ser ampliada a outras áreas da filosofia, como, por exemplo, Ladyman e Ross (2007) propuseram à metafísica e Kornblith (2021) à epistemologia.

⁶ Kim Sterelny (2012), um dos principais filósofos contemporâneos da mente e da biologia, também expressa o endosso dessa metodologia.

⁷ O outro livro examinado por Schliesser foi o “*From Bacteria to Bach and Back*” de Daniel Dennett (2017), outro filósofo que exemplifica bem o espírito naturalista.

1 O QUE É A CONSCIÊNCIA?

“Consciência” é um termo ambíguo ou, como diria Ned Block (1995), um conceito “híbrido”. Os seus mais variados usos são feitos no cotidiano, suscitando confusões. No entanto, essas confusões não se restringem aos usos corriqueiros, pois estão presentes também no uso feito por pesquisadores das ciências cognitiva, neurociência e filosofia da consciência. Neste capítulo, buscarei realizar as distinções necessárias, a fim de tentar elucidar o uso que farei do termo “consciência”, especificamente em seu sentido fenomenal. Em sequência, abordarei o debate atual sobre a possível distinção entre a consciência fenomenal e a consciência de acesso segundo o argumento do transbordamento. Posteriormente, apresentarei os problemas da distribuição e da mensuração da consciência entre os animais. O primeiro diz respeito à extensão da consciência entre as diferentes linhagens de animais. O segundo trata sobre qual seria o método adequado para indicar a presença da consciência nos animais. Por fim, apresentarei de forma panorâmica as três abordagens que buscam solucionar tais problemas: a abordagem teórica, a epistêmica e a da percepção direta.

1.1 Explicitando a Consciência

Considere as seguintes frases do nosso dia a dia:

- a) Será que ele tem consciência desta questão?
- b) Ela sofreu um acidente e perdeu a consciência.
- c) Ele retomou a consciência depois de um longo e pesado sono de descanso.

Como pode-se observar, em cada uma delas o termo “consciência” é utilizado de uma forma diferente. No entanto, não são apenas os nossos usos corriqueiros do termo que estão permeadas de confusões, mas também o seu próprio uso teórico no empreendimento científico e filosófico (CARRUTHERS, 2019; HARLEY, 2021). A mal aplicabilidade do termo pode acarretar a proliferação de diversos mal-entendidos, o que produziria consideráveis erros teóricos. Logo, é de crucial importância diferenciarmos os variados sentidos de “consciência” para que ela seja utilizada de forma clara ao longo do texto.

Aqui o alvo será a chamada *consciência fenomenal* (*phenomenal consciousness*) (BLOCK, 1995). Diferentemente dos conceitos de “triângulo” e “quadrado”, não conseguimos definir a consciência fenomenal de forma objetiva. Podemos definir um triângulo, por exemplo, como um polígono que contém três ângulos internos e três lados. Por outro lado, a consciência

fenomenal não permite este tipo de definição por justamente ser um estado mental cuja característica fundamental é a experiência subjetiva. Toda tentativa de definir a consciência fenomenal se verá as voltas com contraexemplos. Considere o seguinte exemplo: “a consciência fenomenal é um estado cerebral com um certo tipo de estrutura química como a nossa”. Pode-se claramente ser concebido um alienígena com uma estrutura cerebral completamente diferente, mas que mesmo assim possui consciência fenomenal. Definições feitas da poltrona, ou de forma *a priori*, não exprimem aquilo que se entende por consciência fenomenal, pois não captam a sua perspectiva de primeira-pessoa (TYE, 2017). Por essa razão muitos dão preferência a locuções da forma “há algo para alguém que é como ser *x*”, “a característica qualitativa”, entre outras (BLOCK, 1995; CARRUTHERS, 2019; HILL, 2010; NAGEL, 1974). Seguindo a literatura, o melhor a se fazer é abandonar qualquer tentativa de fornecer definições. Assim, não apresentarei uma definição, pois não darei condições suficientes e necessárias, mas explicitarei a “consciência” através de exemplos.

Muitos dos filósofos e cientistas utilizam exemplos como forma de explicitar a consciência fenomenal, mas foi provavelmente com Schwitzgebel (2016) que essa estratégia se tornou mais clara (BLOCK, 1995; GAZZANIGA, 2018; SEARLE, 2010; TYE, 2017). Como forma de obter sucesso, a estratégia exige a satisfação de certos requisitos. O principal deles é que tenhamos apenas um único e óbvio conceito ou categoria que se encaixe com os exemplos fornecidos (SCHWITZGEBEL, 2016, p. 227). Quando mostro, por exemplo, imagens de triângulo para um determinado indivíduo, ele deve perceber que claramente elas se referem ao conceito de triângulo e não a de uma outra forma geométrica.⁸ Por sua vez, um exemplo é ruim quando dá margem a ambiguidades. Se um exemplo tiver como referente algum outro candidato que não seja a consciência fenomenal, ele poderá ser considerado ruim.

Como forma de clarificar o conceito, deve-se apresentar não apenas exemplos positivos daquilo que seria a consciência fenomenal, mas também os negativos. Schwitzgebel apresenta uma lista não exaustiva de classificações possíveis de exemplos positivos. São eles: as experiências sensoriais e somáticas, as imagens conscientes, as experiências emocionais, o pensamento e o desejo, as experiências dos sonhos e as experiências sobre as outras pessoas (SCHWITZGEBEL, 2016, pp. 227-229). Todavia, deve-se notar que tais exemplificações cumprem um papel explanatório sem qualquer pretensão metafísica. Tais classificações não

⁸ Poder-se-ia argumentar que não é óbvio que as imagens apresentadas dos triângulos se referem justamente a “triângulo”, mas às linhas, cores, fundo em branco ou qualquer outra propriedade existente nas imagens. Contudo, isso é falho. Ora, a menos que estejamos tratando de casos raros, me parece ser bastante intuitivo que as mentes das pessoas não funcionam desta maneira (SCHWITZGEBEL, 2016, p. 227).

intentam versar sobre a constituição metafísica da consciência. Ademais, os exemplos que veremos a seguir podem se encaixar em mais de uma categoria selecionada, o que não é um problema, pois o seu propósito é unicamente o de tornar claro aquilo pelo qual entenderei por “consciência”. Vejamos alguns dos exemplos.

Ao ler este texto, você está realizando uma experiência visual. Caso olhe a paisagem externa da sua casa pela janela, você também está realizando uma experiência visual. Ao escutar o som do ambiente dentro do seu quarto, você está realizando uma experiência auditiva. Caso esteja pressionando alguma tecla do computador, você provavelmente está realizando uma experiência somática. Ao sentir sede e beber um copo d’água, você provavelmente experienciará prazer. A experiência da dor de cabeça, ou a ardência devido a uma queimadura, também são exemplos de qualidade desagradável. Todos esses são exemplos de *experiências sensoriais e somáticas*.⁹

Agora feche os olhos e tente imaginar as pirâmides do Egito, a floresta amazônica ou o som de seus familiares cantando feliz aniversário. Peço também que se imagine dançando em uma festa, ou correndo na praia, ou mesmo entrando em uma piscina durante um dia quente de verão. Imagine-se esticando o braço para pegar um objeto que está no alto de uma estante. Todos esses são exemplos de *imagens conscientes*.

Todos já tivemos experiências envolvendo o medo, como o de levar um choque durante o banho devido uma corrente elétrica que escapou da resistência do seu chuveiro. Você também provavelmente já se surpreendeu como com uma festa surpresa. Todos já sentimos indignação ao presenciar uma injustiça contra terceiros. Você provavelmente já sentiu alívio ao resolver um problema difícil. Todos esses são exemplos de *experiências emocionais*.

Todos já desejamos um sorvete em um dia quente ou um chocolate como sobremesa a noite. Quem nunca se perguntou em tom de surpresa: “como eu fiz aquilo?” Ou pensou: “como eu gostaria de ter uma casa naquele bairro”. Você provavelmente já desejou que a chuva passasse para que pudesse ir a um determinado local. Estes são exemplos de *pensamentos e desejos fenomenalmente conscientes*.

Apesar dos sonhos não serem considerados consciência *de algo*, Schwitzgebel considera que os sonhos são, de acordo com a psicologia científica e a psicologia popular (*folk psychology*), fenômenos pertencentes à consciência fenomenal (SCHWITZGEBEL, 2016, pp. 228-229). Quando sonhamos, estamos claramente passando por uma experiência subjetiva (HARLEY, 2021). Pense, por exemplo, no seu último sonho vívido que lhe acordou

⁹ A menos que você seja portador de alguma deficiência visual, auditiva ou de outra natureza sensorial.

repentinamente com a respiração ofegante, ou no sonho em que você estava comendo uma deliciosa refeição. Ambos são exemplos de *experiências de sonho*.

Por fim, temos as experiências envolvendo *outras pessoas*. Normalmente assumimos que as outras pessoas possuem todas ou quase todas as experiências conscientes que elenquei acima. Assim, as experiências que envolvem as outras pessoas também podem ser consideradas exemplos positivos de consciência fenomenal.

Ademais, também existem os exemplos negativos do que não se considera consciência fenomenal. São alguns deles: o crescimento das unhas e dos pelos, a digestão e a secreção hormonal. Não há consciência fenomenal visual do que está acontecendo fora do seu campo de visual. Não temos consciência fenomenal de tudo aquilo que ocorre ao nosso redor imediato. Tampouco há consciência fenomenal de imagens que estão escondidas durante os experimentos de mascaramento (DEHAENE et al, 2001; SCHWITZGEBEL, 2016, p. 229).¹⁰

Devo salientar que em nenhum momento defendi que os exemplos supracitados se referem a uma propriedade intrínseca ou relacional da experiência, ou ainda a uma propriedade física ou não física. Tampouco me comprometi com a tese de que a consciência fenomenal seja algo inefável, privado, irreduzível e de impossível acesso através da investigação científica. Em suma, não assumi até aqui nenhum comprometimento metafísico com o que seria a natureza última da consciência fenomenal. Como Schwitzgebel (2016), o único compromisso aqui feito é com a existência de um sentido pré-teórico de consciência presente na psicologia popular.

Se porventura o leitor ainda tiver dúvidas quanto ao que se entende aqui por “consciência fenomenal”, apenas poderei fornecer mais exemplos intuitivos. Como supostamente Louis Armstrong disse quando perguntado sobre o que seria o jazz, poderia dizer o mesmo quanto à consciência fenomenal: “Cara, se você tem de perguntar, você nunca saberá!”¹¹ (TYE, 2017, p. 7). Em resumo, devemos resistir a todo custo à tentação de defini-la. Considere a seguinte ilustração. Imagine que nos é incumbida a tarefa de ensinar o conceito “azul” para um indivíduo cuja língua materna não possui um termo correspondente. O que se pode fazer é mostrar diversas imagens e objetos de cor azul, como uma caneta, uma pintura e a foto do céu ou do mar, em diversos contextos e tonalidades de cor azul. A cada vez que um objeto é apresentado, é dito: “isto é azul”. Além disso, apresentam-se os mesmos objetos, mas

¹⁰ Dentre os diversos tipos de experimentos de contraste mínimo, os de mascaramento são um dos mais célebres. Grosso modo, ele consiste em tornar uma figura alvo invisível por meio da sua projeção, em um tempo t , no meio de uma sequência de formas de imagens similares que agem como *mascaradores* da figura alvo. Cf. Dehaene, 2014, pp. 37-41.

¹¹ “*Man, if you gotta ask, you’ll never know!*”. Os trechos com tradução minha são os que possuem a versão original na nota de rodapé.

agora com diferentes cores e tonalidades, como vermelho, laranja, preto e branco, e diz “isto não é azul”. Caso a pessoa não seja portadora de alguma deficiência em seu sistema visual, seria implausível que ela não compreendesse o “azul”. Diante disso, seria tentador explicarmos a constituição metafísica do que seria o azul, como, por exemplo, uma frequência de luz que é refletida de um certo modo. Porém, insistir com as apresentações de exemplos ainda seria a forma mais instrutiva para que o indivíduo apreenda o conceito “azul”. O mesmo poderia ser dito da consciência fenomenal (SCHWITZGEBEL, 2016, pp. 234-235).

É importante salientar que tal elucidação com base em exemplos positivos e negativos não implica no abandono das pesquisas científicas e metafísicas sobre a natureza da consciência fenomenal. Mais do que isso, ela contribui com a ciência da consciência na medida em que ajuda na clarificação do termo. Em suma, a consciência fenomenal, cujos exemplos ilustram e captam bem, se refere tão somente a essa natureza experiencial da mente (ANDREWS, 2020a). Claro que, *mutatis mutandis*, estes exemplos positivos e negativos nos permitem estender o que entendemos por “consciência” para os outros animais. Afinal de contas, não existiria uma consciência humana e outra animal, uma vez que os humanos são animais.

Uma outra distinção importante é entre a consciência e a chamada “*consciência de criatura*” (*creature consciousness*). Esta última se apresenta sob duas formas: a *transitiva* e a *intransitiva* (CARRUTHERS, 2019; ROSENTHAL, 2005; VAN GULICK, 2018). Dizemos que um animal, humano ou não humano, possui consciência de criatura transitiva se ele tem ciência (*awareness*), ou percepção de algum objeto, evento externo, ou de seu próprio corpo (CARRUTHERS, 2018a, 2019). A consciência de criatura transitiva também se caracteriza por ser consciência de algo (*conscious of*) (TYE, 2017). Uma capivara, por exemplo, quando percebe um movimento estranho na folhagem da selva, possui consciência desse tipo. Por sua vez, a consciência de criatura intransitiva se identifica com “[...] o fato de estarmos acordados ao invés de dormindo, ou consciente em oposição a um estado comatoso.”¹² (CARRUTHERS, 2019, p. 5). Como Carruthers assinala, quando nos referimos a um estado de consciência de criatura intransitiva em humanos, supõem-se que o indivíduo se encontra em algum outro tipo de estado mental consciente, uma vez que estar acordado normalmente envolve uma experiência consciente de um certo tipo (CARRUTHERS, 2018a, 2019).

¹² “[...] is a matter of being awake rather than asleep, or conscious as opposed to comatose.”

De todas as distinções, a mais controversa e não menos importante é aquela proposta por Ned Block (1995) entre a consciência fenomenal e a de acesso.¹³ A consciência de acesso se caracteriza como aquela cujo conteúdo da representação se encontra disponível ao controle direto, à ação racional, o relato verbal etc. Aqui surge um problema fundamental. Seria a distinção meramente conceitual ou seria também uma distinção empírica? Block sustenta que se trata de uma distinção empírica e fornece argumentos empíricos nos quais a consciência fenomenal não seria acompanhada da consciência de acesso. Segundo ele, a existência de uma é *constitutivamente* independente da outra. Em apoio a sua posição, Block propõe o “argumento do transbordamento” (*the overflow argument*), segundo o qual o conteúdo da consciência fenomenal excede, ou *transborda*, a da consciência de acesso (BLOCK, 1995, 2007a, 2007b, 2008, 2011a, 2014). Vejamos com maiores detalhes na seção a seguir.¹⁴

1.2 O Debate do Transbordamento (*The Overflow Debate*)

Em seu clássico artigo “*On a Confusion About a Function of Consciousness*” de 1995, Block distinguiu a consciência fenomenal da consciência de acesso.¹⁵ Como vimos anteriormente, a consciência fenomenal é aquela caracterizada pela locução “há algo que é tal como se encontrar em um certo estado para alguém” (*what it’s like to be something*) de Nagel (1974). É esse o sentido de consciência que os filósofos têm em mente quando apresentam os seus argumentos modais, como a concebilidade do zumbi filosófico e o quarto da cientista Mary, e não a consciência de acesso que pode ser entendida em termos meramente funcionais (CHALMERS, 1996; JACKSON, 1986). Ou seja, é a consciência fenomenal e não a de acesso que dá ensejo ao chamado “problema difícil” (*hard problem*), o problema da lacuna

¹³ Outros autores realizam mais distinções, como a autoconsciência e a consciência introspectiva. Essas não serão necessárias para o objetivo deste texto. Cf. Hill, 2010; Van Gulick, 2018.

¹⁴ Nesta seção, deixei de lado o debate sobre se a consciência fenomenal é um fenômeno tudo-ou-nada ou gradual, pois ele será abordado no próximo capítulo. As posições tudo-ou-nada, por vezes chamadas de liga-e-desliga, sustentam que a consciência fenomenal é um fenômeno que se está ou não experienciando. Ou seja, ela não admite gradações (CARRUTHERS, 2000, 2018a, 2019; SEARLE, 2006). O gradualismo, por sua vez, admite algum tipo de gradação na experiência fenomenal. Essa posição se apresenta sob duas formas. A primeira é o modelo “*scala naturae*”, onde a variação se dá entre mais e menos consciente em uma escala hierárquica determinada por um critério *C*. Por exemplo, através da quantidade de informação integrada no cérebro, ou Φ^{max} , como sugere a Teoria da Informação Integrada. Segundo essa teoria, quanto maior é a capacidade do sistema de integrar informação, maior é o seu nível de consciência (KOCH, 2020; TONONI e KOCH, 2015). O segundo modelo é o multidimensional, no qual a consciência possui diferentes dimensões que variam formando um dado perfil. Nesta concepção, não existiria um animal com mais ou menos consciência, mas perfis distintos que devem ser analisados de acordo com os diferentes graus de suas dimensões (BAYNE et al, 2016; BIRCH, et al, 2020; FAZEKAS e NEMETH, 2020; FAZEKAS e OVERGAARD, 2016; VEIT, 2022b). Cf. o apêndice para uma representação gráfica dos modelos *scala naturae* e multidimensional da consciência.

¹⁵ A terminologia utilizada por Block (2007a, 2007b, 2008) varia entre consciência de acesso, acessibilidade cognitiva (*cognitive accessibility*) e acesso cognitivo (*cognitive access*). Adotarei o termo “consciência de acesso”.

explanatória e ao argumento do conhecimento antifísicista (CHALMERS, 1996; JACKSON, 1986; LEVINE, 1983). Mas no que consistem tais argumentos? Grosso modo, um zumbi filosófico é um indivíduo que possui as mesmas características físicas e os mesmos engajamentos comportamentais que um ser humano típico, mas que não possui estados mentais conscientes fenomenais. O argumento da concebibilidade, como o do zumbi, pode ser expresso da seguinte forma, tal como Chalmers (2010b, p. 108) o faz. Considerando P a conjunção de todas as verdades da microfísica de nosso universo e Q uma dada verdade fenomenal, como “ x é consciente”, segue-se que:

- A) É concebível que $P \wedge \neg Q$.
 - B) Se é concebível que $P \wedge \neg Q$, então é metafisicamente possível que $P \wedge \neg Q$.
 - C) Se é metafisicamente possível que $P \wedge \neg Q$, então o materialismo é falso.
-
- D) O materialismo é falso.

O argumento do quarto da cientista Mary, por sua vez, diz respeito a uma mulher que viveu desde sempre trancada em um quarto onde somente tinha contato com as cores preto e branco. Com efeito, a cientista passou a sua vida inteira estudando todos os aspectos físicos relacionados às cores. Ela conheceria absolutamente tudo sobre o assunto, inclusive as coisas que a nossa ciência atual não compreende. Em um certo dia, Mary consegue sair de seu quarto e por meio da visualização de uma maçã, ela percebe uma nova cor, a saber, o vermelho. A seguinte pergunta é suscitada: afinal, Mary conseguiu ou não obter um novo tipo de conhecimento? Caso a resposta fosse sim, isso poderia implicar que, por mais que ela tivesse todo o conhecimento físico acerca das cores, existiriam propriedades não físicas e infáveis da experiência subjetiva, os chamados *qualia* (JACKSON, 1986).¹⁶ Chalmers (2010b, p. 109) expõe o argumento da seguinte forma:

- A) Existem verdades sobre a consciência que não são deduzíveis de verdades físicas.
- B) Se existem verdades sobre a consciência que não são deduzíveis de verdades físicas, então o materialismo é falso.

¹⁶ Alguns autores argumentam que não seria necessário o endosso de um dualismo para ser um realista acerca dos *qualia* ou qualquer outra propriedade intrínseca da experiência (BLOCK, 1996; CHALMERS, 2017).

C) O materialismo é falso.

O problema difícil, por seu turno, diz respeito a como e por que os processos físicos, como os realizados pelo nosso cérebro, são acompanhados de experiências qualitativas (CHALMERS, 1995, 1996). O problema difícil se diferencia de outros dois problemas, o “fácil” e o “real”. O problema fácil se refere a como o cérebro é capaz de integrar informação, discriminar estímulos do ambiente, gerar comportamentos flexíveis, produzir propriedades funcionais entre outros. Ele recebe esse nome não porque a sua resolução é fácil, mas em razão dos métodos padrões das ciências cognitivas e das neurociências conseguirem prover um tratamento adequado ao problema (CHALMERS, 1995). Já o problema real, grosso modo, expressa que assumindo a existência da consciência os objetivos mais importantes da ciência da consciência devem ser explicar, prever e controlar as propriedades fenomenológicas que a experiência consciente possui (SETH, 2021, pp. 25-28). Este seria um problema de natureza metodológica cujas ciências e filosofia da consciência devem enfrentar.

Por fim, os Argumentos do Conhecimento se assentam no descompasso que supostamente existiria entre as explicações fisicistas, em terceira-pessoa, das experiências qualitativas, que seriam em primeira-pessoa. Haveria, assim, uma lacuna explanatória entre esse modo de explicação e a natureza dos estados mentais conscientes (LEVINE, 1983). Chalmers (2010b, p. 110) expõe da seguinte maneira esse argumento:

- A) Existe uma lacuna explanatória entre verdades físicas e fenomenais.
- B) Se existe uma lacuna explanatória entre verdade físicas e fenomenais, então existe uma lacuna ontológica, e o materialismo é falso.

C) O materialismo é falso.

Na seção anterior, o conceito de “consciência fenomenal” foi explicitado através da locução de Nagel (1974) e por meio do uso de exemplos (SCHWITZGEBEL, 2016). Em contrapartida, definiu-se a consciência de acesso como aquela cujo conteúdo se encontra disponível aos diversos sistemas consumidores, como a de tomada de decisão e do relato verbal. A minha crença de que Brasília é a capital do Brasil pode ser um exemplo de consciência de

acesso, na medida em que a informação está disponível para o uso de sistemas como o do relato verbal e da racionalidade (BLOCK, 1995, 2007a, 2007b, 2008, 2011a).¹⁷

Carruthers sustenta que existem dois modos de interpretação do conceito de consciência de acesso: uma *disposicional* e outra *categorial* (CARRUTHERS, 2017a, pp. 65-66). Segundo ele, Block inicialmente entendia a consciência de acesso pela primeira forma (BLOCK, 1995), mas posteriormente passou a utilizar a segunda (BLOCK, 2011a). Um conteúdo é interpretado como da consciência de acesso pelo modo disposicional quando se torna disponível *contrafactualmente* aos variados sistemas consumidores. Ou seja, ela se encontra disponível, mas não está necessariamente sendo globalmente compartilhada. Por outro lado, o modo categorial a entende como a própria informação que está sendo globalmente compartilhada. Ela é, portanto, a própria informação transmitida aos sistemas consumidores (CARRUTHERS, 2017a). A partir deste momento, entenderei a consciência de acesso na forma categorial.

Como foi observado ao final da seção anterior, há um extenso debate acerca da distinção entre consciência fenomenal e de acesso. Alguns defendem que, embora sejam conceitualmente distintos, empiricamente elas se correlatam (CHALMERS, 1997). Outros sustentam que a própria distinção conceitual seria falha e que os dois conceitos se referem a um mesmo fenômeno (COHEN e DENNETT, 2011; COHEN, DENNETT e KANWISHER, 2016; CARRUTHERS, 2018a, 2019, 2020a; KOUIDER et al, 2011; OVERGAARD, 2018). Por fim, existem aqueles, mais notadamente Block, que sustentam que cada conceito diz respeito a um fenômeno empírico distinto (BLOCK, 2007a, 2007b, 2008, 2011a, 2014; BRONFMAN et al, 2019).¹⁸

O argumento mais central que sustenta a última posição é a do chamado “transbordamento” (*overflow*) (BLOCK, 2007a, 2007b, 2008, 2011a, 2014).¹⁹ Será dele que me ocuparei nesta seção. Segundo o argumento do transbordamento, devemos realizar tal distinção, pois a consciência fenomenal, por ser mais rica, *transborda* a capacidade limitada da consciência de acesso. Ou seja, o conteúdo da consciência fenomenal excede a capacidade da consciência de acesso. As duas propriedades seriam conceitualmente e empiricamente distintas, com a primeira se referindo aos aspectos qualitativos da experiência e a outra aos aspectos

¹⁷ Existe um debate se os próprios processos cognitivos possuem intrinsecamente uma natureza fenomenal, a chamada “fenomenologia cognitiva”. No entanto, não a abordarei nesta dissertação. Para maiores informações, cf. Bayne e Montague, 2011.

¹⁸ Não possuo o intuito de exaurir a literatura que toca neste debate, pois fugiria do escopo desta dissertação.

¹⁹ O filósofo Matthias Michel (2020) sustenta que esse tipo de argumento e debate já estava presente no século XIX entre os filósofos e “cientistas da consciência” da época.

cognitivos e funcionais. Os mecanismos neurais subjacentes da consciência fenomenal e de acesso seriam distintos, o que poderia ser observado através de experimentos empíricos. Block embasa os seus argumentos especialmente nos experimentos psicológicos baseados no paradigma de Sperling e nos experimentos neurocientíficos realizados pelos neurocientistas da Universidade de Amsterdã (SPERLING, 1960; LANDMAN et al, 2003; SLIGTE et al, 2009).²⁰

No experimento de Sperling, apresenta-se durante um curto período de tempo aos participantes um arranjo de 3 fileiras com 4 letras cada (Figura 2). Logo após, indaga-se quantas letras eles se recordavam. Apesar dos participantes afirmarem terem visto todas ou quase todas as letras, eles somente conseguiam relatar entre 3 e 4 do conjunto total de 12. Em um segundo momento do experimento, após a projeção do arranjo de letras, era exibido um sinal (*cue*) em forma de 3 timbres de sons: um alto, um médio e, por último, um baixo. Quando o sinal (*cue*) apresentado era o de timbre alto, o participante deveria relatar a primeira fileira do arranjo. No tom médio, a fileira do meio. E no tom baixo a última fileira. Os participantes neste momento eram capazes de relatarem quase todas as letras dispostas nas fileiras requisitadas pelos seus respectivos sinais sonoros (*cues*) (SPERLING, 1960).

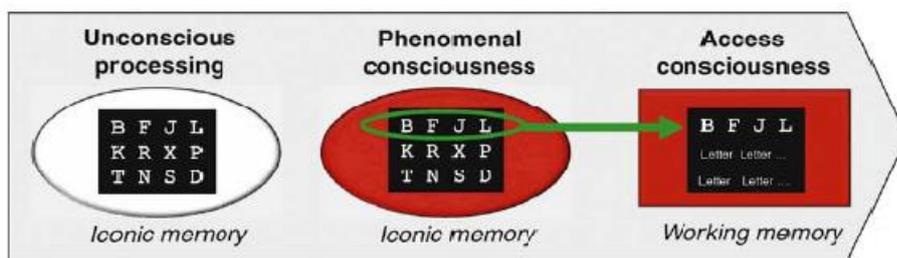


Figura 2 - a) Processamento inconsciente; Memória Icônica, (b) Consciência fenomenal; Memória Icônica, (c) Consciência de acesso; Memória de trabalho; Letra. Experimento de Sperling interpretado por Block. As cores representam o nível de confiança dos participantes na visualização do estímulo, suas gradações vão do nulo (branco) ao alto (vermelho). Kouider et al, 2010, p. 302. Adaptado.

Block interpreta esse experimento como uma evidência importante e favorável à tese do transbordamento (BLOCK, 2007a, 2007b, 2008, 2011a;). Segundo ele, os indivíduos conseguiriam representar conscientemente todas, ou quase todas, as letras do arranjo. No entanto, somente 3 ou 4 letras eram relatadas devido à capacidade limitada da consciência de acesso. Deste modo, o conteúdo da consciência fenomenal seria muito maior em comparação com a de acesso. Importante ressaltar que este tipo de argumento não considera que tais itens são inacessíveis à consciência de acesso, mas apenas que a maior parte deles não podem ser

²⁰ Aqui me delimitarei a abordar apenas o experimento original de Sperling (1960).

acessados imediatamente devido a sua capacidade limitada (BLOCK, 2011a). Block faz a seguinte analogia: nem todos podem ganhar na loteria, no entanto, isto não quer dizer que a loteria é impossível de ser ganha (BLOCK, 2011a, p. 567). É necessário salientar que Block e os proponentes do argumento do transbordamento identificam a consciência de acesso com um tipo específico de memória, a saber, a memória de trabalho (BLOCK, 2007a, 2007b, 2008, 2011a; BRONFMAN et al, 2014; BRONFMAN et al, 2019). A memória de trabalho, grosso modo, é um mecanismo mental complexo que sustenta, manipula e processa informações por um curto período, sendo responsável por contribuir e guiar comportamentos flexíveis, como as de tomada de decisão.²¹ Ela é classificada por muitos pesquisadores como um tipo de memória de curto prazo (ABEN et al, 2012; BADDELEY e HITCH, 1974; CARRUTHERS, 2013; CHAI et al, 2018; COWAN, 2008; SHEVLIN, 2020). Há relativo consenso quanto à sua capacidade de armazenagem sendo de 3 a 4 itens (COWAN, 2000, 2010). Ademais, a sua base neuronal geralmente é identificada no córtex pré-frontal, especificamente na região dorsolateral (Figura 3) (CHAI et al, 2018; CURTIS e D'ESPOSITO, 2003; FINN et al, 2019; FUNASHI, 2017; GOLDMAN-RAKIC, 1995; KIM et al, 2015).²² Outras evidências sugerem que áreas como a do lobo temporal medial também estão envolvidas com a memória de trabalho (BORAN et al, 2022; KORNBLITH et al, 2017).

O córtex pré-frontal (CPF) é uma região que se encontra no lobo frontal do cérebro e que cobre boa parte do córtex. Ele pode ser dividido em 5 grandes áreas: o CPF anterior, o CPF caudal, o CPF ventrolateral, o CPF dorsolateral e o CPF medial (Figura 3) (KANDEL et al, 2021; MICHEL e MORALES, 2020, p. 494).²³ Apesar do CPF não ser a área responsável por gerar as informações sensoriais, ele as recebe por meio das regiões que são responsáveis, as processa e, em virtude de suas extensas conexões, distribui globalmente às outras regiões do cérebro. O CPF é comumente associado às funções complexas²⁴ de controle e monitoramento, como as de planejamento, tomada de decisão, sociabilidade, processamento avaliativo dos estímulos afetivos de risco e recompensa, entre outras (KANDEL et al, 2014; KANDEL et al, 2021; MICHEL e MORALES, 2020; PASSINGHAM, 2021; WARD, 2020). Dependendo do hemisfério abordado, é possível que determinadas regiões do CPF possuam funções diferentes.

²¹ Para maiores informações sobre a memória de trabalho, cf. Logie et al, 2021. Sobre a possível centralidade da memória de trabalho na mente e na consciência humana, cf. Carruthers, 2015.

²² Sobre o córtex pré-frontal, cf. Passingham e Wise, 2012; Passingham, 2021.

²³ Existem também as sub-regiões dessas áreas, como o CPF frontopolar e o CPF orbitofrontal. Este último se localiza em parte na região do CPF anterior e outra no CPF ventrolateral. Cf. Michel e Morales, 2020; Passingham e Wise, 2012; Passingham, 2021.

²⁴ Alguns dizem que o CPF é a região central das chamadas “funções executivas” (SAPOLSKY, 2017; WARD, 2020). Richard Passingham (2021, p. 8), entretanto, considera vago o uso desse termo.

O hemisfério esquerdo do CPF ventrolateral, por exemplo, talvez possua a função de “[r]ecuperar e manter informações semânticas e/ou linguísticas ([á]reas 44 + 45 na esquerda também chamada área de Broca)”. Em contrapartida, tudo parece indicar que o hemisfério direito possua a função de “[r]ecuperar e manter informações visuais e/ou espaciais”²⁵ (WARD, 2020, p. 388). Muitos pesquisadores defendem que o CPF é a região responsável, ou a que mais contribui, para a consciência (BAARS, 1989; BROWN et al, 2019; DEHAENE e CHANGEUX, 2011; DEHAENE, 2014; LEDOUX, 2019; MASHOUR et al, 2020; MICHEL e MORALES, 2020).²⁶

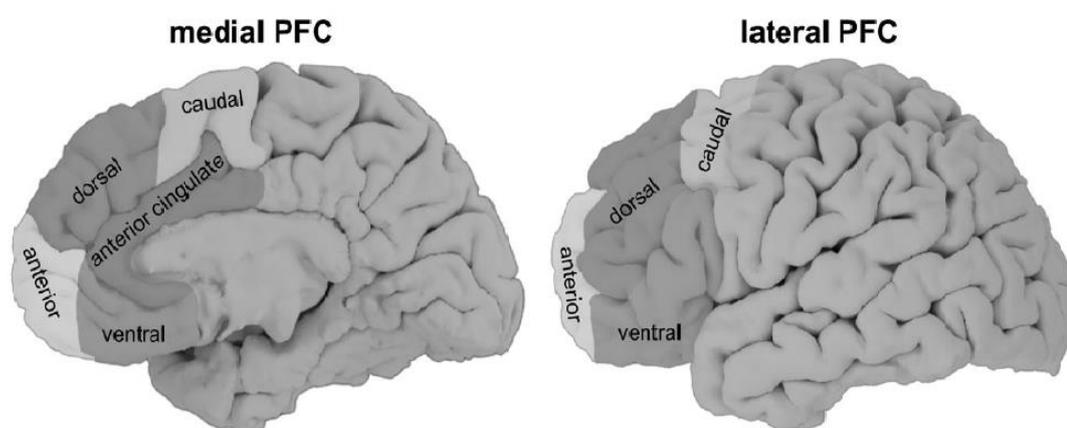


Figura 3 – O córtex pré-frontal (PFC, sigla em inglês para *Prefrontal Cortex*) e as suas regiões. Michel e Morales, 2020, p. 495.

Block, por outro lado, sustenta que o experimento de Sperling envolve dois tipos de memórias. Segundo ele, os conteúdos de ambas são mais ricos que a memória de trabalho. São elas: a memória icônica e a memória visual frágil de curto prazo (MVFCP) (BLOCK, 2011a). A primeira é caracterizada como uma memória de curto prazo, na qual as representações visuais persistem por um período extremamente curto de tempo após o

²⁵ “*Retrieval and maintenance of semantic and/or linguistic information (Areas 44 + 45 on left also called Broca’s area)*”; “*Retrieval and maintenance of visual and/or spatial information*”.

²⁶ Há um grande debate acerca da base neuronal da consciência humana. Atualmente, dois grandes conjuntos de teorias estão em disputa. Existem as teorias que defendem o CPF como a região responsável, ou a que mais contribui, pela consciência e as teorias das regiões posteriores do córtex. Para as primeiras, as teorias mais famosas são as de ordem superior e as do tipo globalista (BAARS et al, 2021; BROWN et al, 2019; DEHAENE, 2014; LAU, 2022; LAU e ROSENTHAL, 2011; MASHOUR et al, 2020). Já a segunda, são as conhecidas como teorias localistas. Estas consideram que são as regiões posteriores do cérebro, como as do lobo occipital, as áreas responsáveis por realizarem a consciência (BLOCK, 2007a, 2011a; LAMME, 2006). Importante notar que algumas teorias, como certas versões da teoria do espaço de trabalho global, reduzem a consciência fenomenal à de acesso (CARRUTHERS, 2019; NACCACHE, 2018). Não entrarei em maiores detalhes sobre este debate, pois fugiria do escopo desta dissertação. Para maiores informações, cf. Bellet et al, 2022; Bartels, 2021; Boly et al, 2017; Kapoor et al, 2022; Malach, 2021, 2022; Michel, 2022; Michel e Morales, 2020; Michel e Doerig, 2021; Naccache et al, 2021; Raccach, Block e Fox, 2021; Weilhhammer et al, 2021.

desaparecimento do estímulo. Esse tipo de memória possui em sua composição aspectos da própria retina do indivíduo. Apesar de sua duração decair rapidamente, durando cerca de 100 milissegundos, ela seria muito mais rica que a memória de trabalho (BLOCK, 2007a, 2011a). Ademais, evidências em macacos sugerem o córtex V1 como a sua base neuronal (TEEUWEN et al, 2021). Por sua vez, a MVFCP consiste em um tipo de memória de curto prazo que possui uma duração maior do que a memória icônica. Embora a MVFCP possua uma riqueza menor do que a memória icônica, ela ainda é significativamente maior do que a da memória de trabalho. Em 1 segundo de projeção do estímulo com sinal (*cue*), a MVFCP consegue armazenar cerca de 7 de 8 itens, em 2.5 segundos 6 de 8, e em 4 segundos 5 de 8 itens. Como visto, muito acima da capacidade de 3 a 4 itens da memória de trabalho (BLOCK, 2011a, p. 571). O córtex V4 (Figura 4) seria a área cerebral responsável pela realização e sustentação da MVFCP, e não as regiões inferiores do córtex visual V1, V2 e V3 (BLOCK, 2011a; SLIGTE et al, 2009). Block, identifica a consciência fenomenal com o conteúdo da MVFCP, localizado nesta região posterior do córtex, em áreas elevadas e médias (*mid-high*) do córtex visual. Por outro lado, o CPF dorsolateral seria responsável pela consciência de acesso, identificada pela memória de trabalho (BLOCK, 2007a, 2011a).

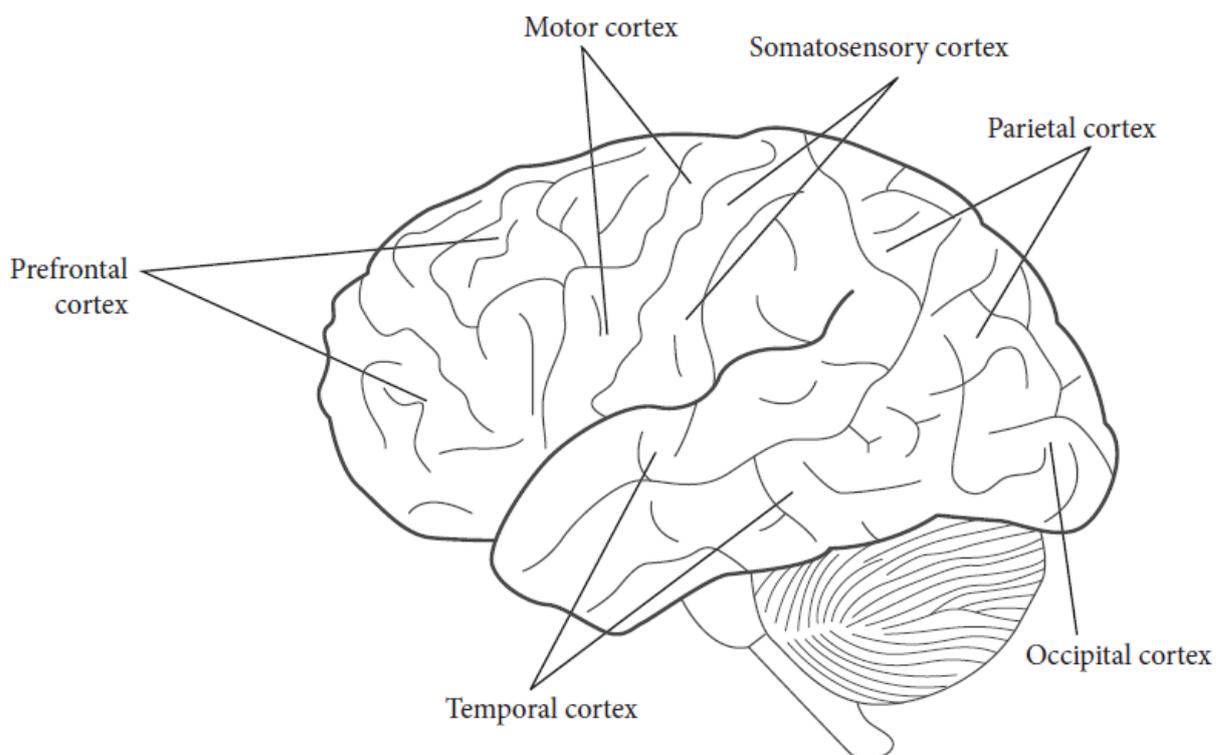


Figura 4 – As principais regiões do cérebro. Hemisfério esquerdo, visão de fora. Tradução da esquerda à direita: Córtex Pré-frontal; Córtex Motor; Córtex Somatossensorial; Córtex Parietal; Córtex Occipital; Córtex Temporal. Carruthers, 2015, In: Figures.

Em suma, o que o experimento de Sperling nos revelaria, de acordo com Block, seria que os participantes possuiriam consciência fenomenal de todas, ou quase todas, as letras do arranjo. No entanto, devido a capacidade limitada da memória de trabalho, tornar-se-ia impossível que todas essas informações entrassem em um mesmo momento na consciência de acesso. Deste modo, por mais que os participantes relatassem que tinham percebido todas, ou quase todas, as letras do arranjo, em razão da capacidade de armazenagem limitada da memória de trabalho, somente uma pequena parte delas eram relatadas. A maior riqueza da consciência fenomenal seria, por fim, comprovada por meio da segunda fase do experimento. Nela, os sujeitos recebiam um sinal (*cue*) após o desaparecimento do estímulo. Neste momento, eles conseguiam relatar quase todas as letras da linha requisitada pelo sinal (*cue*). Isto evidenciaria o transbordamento da consciência fenomenal face à consciência de acesso. Com efeito, a consciência fenomenal seria mais rica (*rich*) em comparação com a consciência de acesso que seria mais escassa (*sparse*). Os resultados dos experimentos corroborariam a tese de que a consciência fenomenal e a consciência de acesso são propriedades distintas que se realizam em mecanismos cerebrais distintos. A consciência fenomenal se realizaria no córtex V4 e a consciência de acesso no CPF dorsolateral (BLOCK, 2007a, 2007b, 2008, 2011a; SPERLING, 1960).

1.2.1 Interpretações Contrárias

Por diferentes vias, inúmeros pesquisadores criticaram o argumento do transbordamento de Block e a sua interpretação do experimento de Sperling (PHILLIPS, 2011, 2018; OVERGAARD, 2018). Vejamos a seguir algumas destas críticas.

No que tange os experimentos baseados no paradigma de Sperling, as abordagens genéricas (*generic phenomenology*) e fragmentadas (*fragmentary phenomenology*) da fenomenologia são duas das mais importantes interpretações contrárias à de Block (COHEN e DENNETT, 2011; COHEN, DENNETT e KANWISHER, 2016; KOUIDER et al, 2010). Segundo a abordagem genérica, antes de receberem o sinal (*cue*), os participantes do experimento armazenam *inconscientemente* somente formas genéricas das letras do arranjo e, por meio da atenção, apenas um pequeno número de letras de forma detalhada. Ou seja, os sujeitos perceberiam algumas letras de forma específica, enquanto todo o restante se constituiria de formas genéricas. Isto se contraporía à interpretação do transbordamento, segundo a qual eram percebidas todas ou quase todas as letras de forma detalhada (BLOCK, 2007a, 2011a; COHEN e DENNETT, 2011; COHEN, DENNETT e KANWISHER, 2016). Após o

fornecimento do sinal (*cue*), os sujeitos conseguiriam relatar as letras requisitadas devido ao acesso a essas representações que estariam armazenadas inconscientemente. Seria a captura dessas representações inconscientes por meio da consciência de acesso que garantiria o *estatuto* consciente de seus conteúdos (COHEN e DENNETT, 2011, p. 359).

Em um segundo artigo, Cohen, Dennett e Kanwisher (2016) sustentam que a aparente maior riqueza entre aquilo que seria percebido e aquilo que é relatado se deve em parte ao *background* estatístico da cena que está diante de si, e não pela capacidade limitada de 4 itens que a memória de trabalho possui, como Block (2011a) sugere. A quantidade de informação ao qual os indivíduos têm acesso seria superior à de 4 itens a cada instante. Ora, disto resulta que os participantes “percebem um montante de itens com uma qualidade elevada de fidelidade, enquanto o restante do mundo é representado como uma semelhança estatística (ou um conjunto de estatísticas).”²⁷ (COHEN, DENNETT e KANWISHER, 2016, p. 332). A percepção seria mais rica do que os proponentes do transbordamento acusam as posições divergentes. A cena é, desta forma, percebida em sua completude com uma certa riqueza. A diferença existente seria apenas entre alguns subconjuntos que seriam percebidos sem atenção e, *a fortiori*, com baixa fidelidade, e outros que seriam percebidos com atenção e, *a fortiori*, com alta fidelidade. Isto se deve em virtude da nossa limitação cognitiva, seja via atenção, ou devido a própria capacidade circunscrita da memória de trabalho. O restante da visão, que está na periferia da atenção, formaria uma semelhança estatística (Figura 5) (COHEN, DENNETT e KANWISHER, 2016). Em resumo, o que os indivíduos percebem em uma determinada cena são apenas alguns itens imediatamente identificados com uma qualidade alta, enquanto na periferia da atenção formar-se-ia um *gist*, uma semelhança estatística com baixa definição. Após o recebimento do sinal (*cue*), as *representações inconscientes* do sujeito seriam acessadas, enriquecidas e tornadas conscientes e, assim, aptas para serem relatadas.

²⁷ “[...] a handful of items are perceived with high fidelity, while the remainder of the world is represented as an ensemble statistic (or set of statistic).”

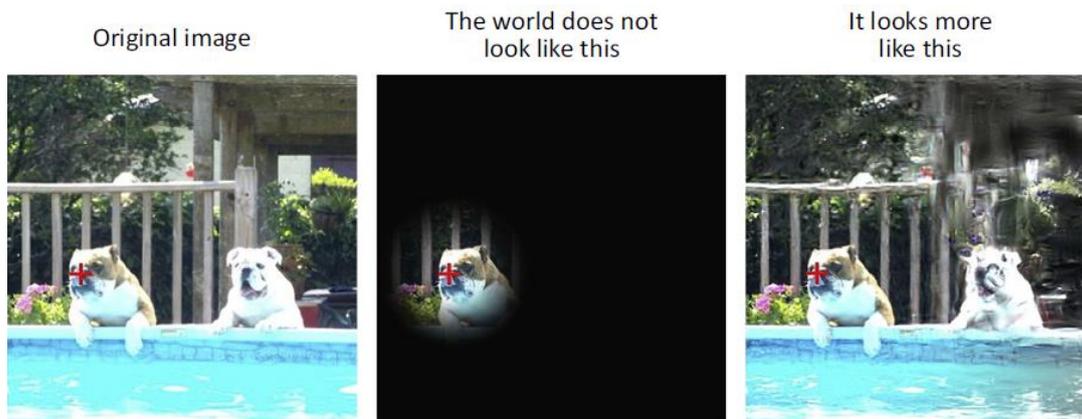


Figura 5 – (a) Imagem original, (b) O mundo não se parece como isto, (c) Se parece mais como isto. A cruz vermelha denota a localização que a atenção está direcionada. A primeira imagem diz respeito a cena original, a segunda como Block (2011a) interpreta os seus adversários e a terceira a interpretação genérica da semelhança estatística. Cohen, Dennett e Kanwisher, 2016, p. 326. Adaptado.

A outra importante concepção contrária ao argumento do transbordamento é a abordagem fragmentada da fenomenologia (KOUIDER et al, 2010). Segundo ela, os participantes do experimento de Sperling possuem acesso consciente a uma alta quantidade de informações com baixo nível de detalhe, que são os fragmentos, e uma quantidade inferior de informações detalhadas (aquelas que se encontram na memória de trabalho). Após fornecido o sinal (*cue*), as representações fragmentadas dos participantes seriam preenchidas, ou reconstruídas, de forma ilusória, tornando a percepção de riqueza uma ilusão construída pela própria cognição dos indivíduos (KOUIDER et al, 2010, p. 302) (Figura 6). Desse modo, a aparente riqueza que os participantes relatam no experimento não seria devido ao transbordamento da consciência fenomenal sob a de acesso, mas em virtude da ilusão gerada pela cognição.

Contudo, Block sustenta que, durante o experimento de Sperling, o relato dos participantes entraria em conflito com as interpretações propostas pelas duas abordagens descritas acima. A razão seria a seguinte: mesmo após os indivíduos afirmarem perceberem entre 3 ou 4 letras, os “participantes – e experimentadores servindo como participantes – continuariam insistindo que estão simultaneamente conscientes de *todos* os elementos no arranjo.”²⁸ (BAARS, 1988, p. 15, *apud* BLOCK, 2011a, p. 570). Isso contradiria as concepções genéricas e fragmentadas da fenomenologia.

²⁸ “[...] subjects – and experimenters serving as subjects – continue to insist that they are momentarily conscious of *all* the elements in the array.”

No entanto, a evidência fornecida por Block (2011a), via comentário de Baars (1988), seria apenas de natureza anedótica, isto é, uma evidência informal em razão da forma com que os relatos foram colhidos. Ela é caracterizada desta forma, devido à falta de um procedimento rigoroso, sistemático e específico que um experimento científico demanda. Para que ela se tornasse uma evidência robusta a favor do argumento do transbordamento, um tratamento científico seria exigido.

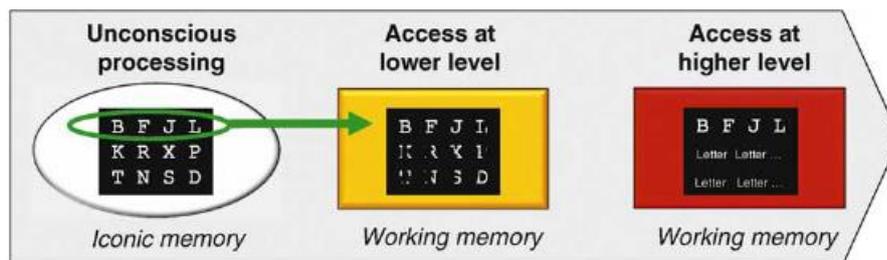


Figura 6 – (a) Processamento inconsciente; Memória icônica, (b) Acesso ao nível inferior; Memória de trabalho, (c) Acesso ao nível elevado; Memória de trabalho; Letra. As cores representam o nível de confiança dos participantes na visualização do estímulo, suas gradações vão do nulo (branco) ao alto (vermelho), passando pelo médio-alto (amarelo). Interpretação do experimento de Sperling segundo a abordagem fragmentada da fenomenologia. Kouider et al, 2010, p. 302. Adaptado.

Com o objetivo de prover justamente um tratamento rigoroso ao conteúdo das experiências relatado pelos participantes, Cova e colegas (2020) reproduziram o experimento de Sperling²⁹ com as devidas modificações. Os pesquisadores dividiram em três blocos a reprodução do experimento de Sperling. Em cada bloco, um arranjo de 12 letras era exposto em uma tela de computador diante do participante. Ao término da projeção do arranjo, os participantes eram requisitados a relatarem o máximo de letras das quais se recordassem. Corroborando o experimento original, as respostas indicavam que os participantes se lembravam de algo em torno de 3 a 4 letras do total de 12. Como forma de recolher rigorosamente o dado sobre a experiência dos participantes, ao final de cada bloco solicitava-se que eles respondessem a dois questionários. Ao final do estudo, também era realizada uma entrevista livre com cada um dos participantes, feitas por dois entrevistadores diferentes. Os questionários consistiam em duas perguntas com respostas de múltipla escolha. A primeira pergunta era a seguinte: “Em geral, você sente como se tivesse visto todas as letras projetadas na tela?”³⁰ (COVA et al, 2020, p. 11). As respostas limitavam-se em: “Sim”, “Não” e “Eu não sei”. A segunda pergunta era a seguinte: “Pense sobre o que você experienciou durante o bloco

²⁹ Os pesquisadores também reproduziram o experimento de Landman e colegas (2003), considerado pertencente ao paradigma de Sperling. Entretanto, como dito anteriormente, me limitarei nesta seção a abordar somente o experimento original de Sperling (1960).

³⁰ “In general, did you feel like you saw all the letters displayed on the screen?”

que acabou de terminar. Quais das próximas afirmações que melhor descrevem o que você experienciou quando as letras foram projetadas na tela?”³¹ (COVA et al, 2020, p. 11). Os participantes possuíam disponíveis seis respostas. As duas primeiras refletiam ao argumento do transbordamento, a terceira à abordagem genérica, a quarta e a quinta à abordagem fragmentada e, por fim, a última resposta era uma forma de controle para saber se os participantes tinham realmente entendido a proposta. As respostas para o primeiro questionário foram majoritariamente o “não”, com uma ressalva para apenas o último bloco (Figura 8). Para o segundo questionário, as maiores porcentagens foram relacionadas às respostas que eram de acordo com as abordagens genéricas e fragmentadas da fenomenologia (Figura 7).

	% of yes/no answers	Confidence
Block 1	Yes: 22	Yes: 70.8 (23.2)
	No: 78	No: 74.6 (22.0)
Block 2	Yes: 34	Yes: 76.4 (18.4)
	No: 64	No: 76.7 (18.0)
Block 3	Yes: 50	Yes: 75.4 (19.5)
	No: 42	No: 81.7 (16.1)

Figura 8 - Respostas relacionadas ao primeiro questionário. "Confidence" é o nível de confiança que os participantes possuíam da própria resposta. Tradução da esquerda à direita: “% de respostas sim/não”, “Bloco”, “Sim”, “Não”, “Confiança”. Cova et al, 2020, p. 13.

	Block 1	Block 2	Block 3
All letters in detail and identified	8% (81.3)	4% (95.0)	6% (88.3)
All letters in detail	8% (82.5)	12% (79.2)	14% (77.1)
All letters, but not in detail	34% (79.1)	36% (79.2)	38% (80.0)
Most letters	36% (82.8)	30% (76.0)	26% (81.2)
Small part of the letters	12% (74.2)	12% (80.8)	14% (82.9)
Nothing	0%	2% (40.0)	2% (10.0)

Figura 7 - Respostas relacionadas ao segundo questionário. Tradução da esquerda à direita: “Bloco”, “Todas as letras em detalhes e identificadas”, “Todas as letras em detalhes”, “Todas as letras, mas não em detalhes”, “Maioria das letras”, “Pequenas partes das letras”, “Nada”. Cova et al, 2020, p. 13.

Em relação às respostas das entrevistas, a maior parte delas estavam de acordo com as abordagens genéricas (62% e 52%) e fragmentadas (12% e 12%) da fenomenologia. Enquanto a menor parte, quando somamos as duas abordagens anteriores, estava em

³¹ “Think about what you experienced during the block that just ended. Which of the following statements best describe what you experienced when the letters were displayed on the screen?”

conformidade com o argumento do transbordamento (16% e 20%) (COVA et al, 2020, p. 13). Os entrevistadores também apresentaram aos participantes duas interpretações científicas sobre aquilo o que eles experienciaram, a fim de obter as respostas que mais se adequavam a elas. Uma dizia respeito à interpretação do argumento do transbordamento e a outra se assemelhava com a disjunção entre as abordagens genéricas e fragmentadas da fenomenologia. As respostas foram majoritariamente positivas à segunda interpretação (72% e 72%) (COVA et al, 2020, pp. 14-15).

Todas as respostas do experimento, corroboraram aquilo o que os proponentes das abordagens genéricas e fragmentadas da fenomenologia defendiam. No que tange o conteúdo fenomenal pré-sinal (*pre-cue*) das experiências dos participantes, a hipótese do transbordamento se demonstrou estar baseada apenas em uma evidência anedótica que, processada por um tratamento rigoroso tal qual fornecido por Cova e colegas (2020), demonstrou estar equivocada. Em linhas gerais, a evidência anedótica consiste em uma evidência informal baseada na história de uma experiência subjetiva de um indivíduo. A utilização de evidência anedótica é tipicamente caracterizada por sua ausência de rigor devido a fatores como: a carência total ou parcial de método, a utilização de histórias subjetivas sem uma coleta de dados adequada, a ausência de um n amostral razoável, a impossibilidade de replicação, ser facilmente suscetível a vieses cognitivos³² entre outros. As evidências anedóticas, portanto, devem ser evitadas no empreendimento científico³³, bem como na filosofia que busca estar empiricamente bem fundamentada.

O argumento do transbordamento suscita uma questão adicional que diz respeito à influência que o sinal (*cue*) pode ter na fenomenologia do sujeito. Os experimentos que utilizam “retro-sinais” (*retro-cues*) dizem que eles podem aumentar a experiência consciente de um dado estímulo (CARRUTHERS, 2020a; SERGENT et al, 2013; THIBAUT et al, 2016). Se o estímulo percebido pelo sujeito estivesse sempre consciente, poderíamos supor que o sinal (*cue*) fornecido após o estímulo fosse inefetivo, dado que o conteúdo estaria o tempo todo consciente (CARRUTHERS, 2019, 2020a). Poderíamos supor o seguinte: a representação da MVFCP estaria armazenada de forma inconsciente no cérebro e, à medida que o sinal (*cue*) fosse dado, o conteúdo seria resgatado pela consciência de acesso, tornando-se assim consciente.

³² Para saber mais sobre como a nossa mente é passível de cair em vieses cognitivos, cf. Kahneman, 2012; Kahneman et al, 2021; Pinker, 2021.

³³ Para saber mais sobre como uma evidência anedótica contribuiu de forma negativa à ciência, especificamente a psicologia comparada, cf. o caso célebre do cavalo Clever Hans (ANDREWS, 2020b; WYNNE e UDELL, 2020).

Argumentos adicionais poderiam ser fornecidos contra o transbordamento, como a interpretação gradualista de Overgaard (2018) e a possível implicação pampsiquista que o transbordamento possui (NACCACHE, 2018). Contudo, não possuo o objetivo de revisar em apenas uma seção toda a complexa e extensa literatura sobre este debate. Como tentei mostrar, existem mais motivos para acreditarmos que a distinção entre a consciência fenomenal e de acesso é meramente conceitual e não real. Acredito que a balança do debate pende contra aqueles que defendem o argumento do transbordamento. Isto posto, não voltarei a esse debate nos próximos capítulos.

1.3 Os Problemas da Distribuição e da Mensuração da Consciência Animal

Realizadas as distinções necessárias sobre o uso do termo “consciência”, podemos finalmente avançar em direção ao tema central desta dissertação, a saber, a consciência animal. Algumas perguntas importantes podem ser feitas, como, por exemplo: Até onde se encontra a consciência na árvore filogenética? Em quais linhagens ou espécies ela está presente? Quais são os animais que a possuem? Seria a consciência um fenômeno restrito aos seres humanos? Essas questões são conhecidas na literatura sob a rubrica do “problema da distribuição da consciência” (ALLEN e TRESTMAN, 2016; BAYNE e SHEA, 2020; HARLEY, 2021; GINSBURG e JABLONKA, 2019, 2022). O problema da distribuição, em resumo, diz respeito a quais são os animais dotados de consciência. Note que essa questão é diferente do problema da origem da consciência, na medida em que o primeiro diz respeito à extensão da consciência para além dos seres humanos³⁴, enquanto o segundo trata sobre a própria origem e história evolutiva da consciência, ou seja, como, onde e porque a consciência veio a surgir em um determinado local e momento histórico (DAMÁSIO, 2019; FEINBERG e MALLATT, 2013, 2017, 2018; GINSBURG e JABLONKA, 2019; GODFREY-SMITH, 2020a; LEDOUX, 2019; MASHOUR e ALKIRE, 2013; VEIT, 2022a). O problema da origem da consciência envolve tipicamente as seguintes perguntas: “Seria a consciência um fenômeno antigo ou apenas surgiu recentemente na árvore filogenética?”, “Quais foram as causas que propiciaram o surgimento da consciência?”, “A consciência surgiu como uma adaptação ou um subproduto de um outro traço?”. O problema da distribuição também se distingue do chamado “problema da consciência mínima”. O problema da consciência mínima diz respeito às questões que envolvem as

³⁴ Aqui estou restringindo o problema da distribuição da consciência à apenas os casos dos animais. Desconsidero os casos *também importantes* de: humanos neurodivergentes, fetos, bebês recém-nascidos, pessoas em estado vegetativo, comatoso, sedado, I. A. e organoides cerebrais. Em relação aos casos de bebês, cf. Ferreira, 2017. Sobre pacientes em estado vegetativo, cf. Shea e Bayne, 2010.

condições mínimas e suficientes para que um certo organismo seja considerado dotado de estados mentais conscientes.³⁵

Dizer quais são os animais que possuem consciência é de maneira alguma uma tarefa trivial. Para isso, seria necessário um método confiável e preciso que conseguisse indicar a presença dela em um determinado animal ou espécie. Como podemos determinar se animais tão diferentes³⁶ e distantes na árvore filogenética, como um chimpanzé e um caranguejo-eremita, possuem consciência? A esses problemas denominarei, acompanhando Browning e Veit (2020), de “problema da mensuração da consciência”.³⁷ O problema da mensuração diz respeito a qual seria o método correto a ser utilizado para indicar se um determinado animal possui ou não consciência. Ou seja, enquanto o problema da distribuição trata da extensão da consciência no reino animal, o da mensuração discute questões metodológicas sobre a presença da consciência entre os animais. Note que ambos os problemas estão interligados, uma vez que para que se determine a extensão da consciência entre os animais, devemos ter métodos capazes de indicar a consciência entre eles. Devo salientar que ao utilizar o termo “mensuração” não intuíto me comprometer com a tese gradualista da consciência, ou com a possibilidade de quantificar o quão consciente é um certo estado mental de um dado animal. No entanto, também não estou endossando de forma tácita a tese tudo-ou-nada da consciência. Utilizo “mensuração” apenas no sentido da capacidade de indicar a presença da consciência em um certo sistema cognitivo. Deste modo, me mantenho neutro em relação a teses estritas sobre a natureza da

³⁵ O problema da consciência mínima também pode se relacionar com o problema da *intencionalidade* mínima, uma vez que para muitos autores a capacidade de representar e de ser consciente estão intimamente entrelaçadas ou consistem essencialmente no mesmo fenômeno (e. g. TYE, 2022). Grosso modo, o problema da intencionalidade mínima diz respeito às condições mínimas e suficientes que um determinado organismo deve satisfazer para que seja considerado um sistema representacional genuíno (SOUZA FILHO, 2018, caps. 3 e 4, 2022). Discute-se também qual seria a forma mais primitiva de estados representacionais e conscientes. Alguns autores argumentam que a percepção é a forma mais primitiva (BURGE, 2022; DRETSKE, 1999), enquanto outros sustentam que são as emoções, estados mentais com valência, as mais primevas (DENTON et al, 2009; GODFREY-SMITH, 2017; VEIT, 2022a, 2022b). Cada uma dessas posições implica em uma distribuição de estados representacionais e conscientes diferentes no reino animal. Por exemplo, Burge (2022, p. 8) afirma que por terem percepção visual os artrópodes possuiriam estados representacionais, ao passo que as lesmas e outros gastrópodes não. Por outro lado, Godfrey-Smith (2020b) considera que apesar de não possuírem sistemas perceptuais complexos, como os dos artrópodes, as lesmas poderiam ter estados mentais conscientes primitivos em razão de terem sistemas afetivos minimamente complexos, algo que os artrópodes não possuiriam. Note que a possibilidade de existirem animais conscientes, mas que não possuem capacidades representacionais e vice-versa, atinge apenas as posições que identificam a capacidade de representação e de ser consciente como um mesmo fenômeno. Aqueles que, como Burge (2022, pp. 6-7), metafisicamente distinguem a capacidade de representação e de ser consciente poderiam admitir a existência de animais que representam, mas não são conscientes e vice-versa.

³⁶ Essas diferenças podem ser desde natureza neurofisiológica e anatômica a até comportamental.

³⁷ Seth e colegas (2008) já tratavam sobre o problema da mensuração, mas em um sentido mais geral. O problema da mensuração, em uma versão ampla, busca entender a possibilidade de identificar se um determinado evento sensorial, motor ou cognitivo é um estado mental consciente, em contraposição a um estado mental inconsciente (independentemente do tipo de sistema cognitivo em questão). Cf. também Seth, 2021, cap. 1; Seth e Bayne, 2022, pp. 10-11.

consciência, como o gradualismo ou o tudo-ou-nada, bem como a capacidade da ciência da consciência quantificar estados mentais conscientes.

Isto posto, note que as perguntas realizadas pelos problemas da distribuição e da mensuração também se diferenciam do chamado “problema metafísico da consciência animal”. Esse problema diz respeito à constituição metafísica do que é ser um animal com um estado consciente do tipo *E*. Em outras palavras, o problema metafísico da consciência animal se pergunta o seguinte: se há algo para um animal que é como ser consciente, o que metafisicamente isto significa? Ou seja, o que representa para um animal ter um estado mental consciente do tipo *E*? Nesta dissertação me restringirei a apenas os problemas da distribuição e da mensuração.

Combinando a taxonomia de Andrews (2020a) com a de Birch (2020a) e realizando algumas alterações, considero que existem ao menos três abordagens possíveis para responder o problema da mensuração: a abordagem teórica, a abordagem epistêmica e a abordagem da percepção direta.³⁸ Antes de tratar de cada uma, devo esclarecer que a taxonomia de Andrews (2020a) é levemente diferente da que estou apresentando aqui. Ao invés da abordagem da percepção direta, Andrews (2020a) sugere a chamada “abordagem biológica”. Segundo Andrews (2020a, pp. 93-95), a abordagem biológica busca preencher a lacuna explanatória e explicar a consciência pela via de sua própria história evolutiva. Seria apresentando os motivos pelos quais a consciência surgiu e evoluiu ao longo da história dos animais, que forneceríamos a sonhada explicação que falta à neurobiologia, isto é, a definitiva e completa redução dos eventos conscientes aos processos biológicos do cérebro. Andrews cita os filósofos Peter Godfrey-Smith, Evan Thompson e a dupla de biólogas Simona Ginsburg e Eva Jablonka³⁹ como exemplos de autores que perseguem esse tipo de abordagem (ANDREWS, 2020a; GODFREY-SMITH, 2019b, 2020a; GINSBURG e JABLONKA, 2019, 2022; THOMPSON, 2007). No entanto, ao menos nos moldes que Andrews apresenta, essa proposta tem como alvo problemas diferentes daqueles que tratarei neste texto. Por meio da noção de função biológica⁴⁰, esta abordagem estaria mais envolvida em um empreendimento reducionista da consciência e com

³⁸ Jonathan Birch (2020a) realiza uma taxonomia razoavelmente similar à de Andrews (2020a). O que Birch chama de abordagem teórica robusta (*theory-heavy approach*) equivaleria à abordagem teórica de Andrews. Considero ambas a abordagem teórica neutra (*theory-neutral approach*) e a abordagem teórica leve (*theory-light approach*) apresentadas por Birch como pertencentes à abordagem epistêmica.

³⁹ Halina e colegas (2022) expressam que a posição denominada “Aprendizagem Associativa Ilimitada” (AAI), sustentada pela dupla Ginsburg e Jablonka (2019, 2022) e outros autores (BRONFMAN et al, 2016), pode ser caracterizada como uma abordagem epistêmica ou teórica de acordo com a interpretação adotada. Browning e Birch (2022, p. 7) listam a AAI como um tipo de abordagem epistêmica. Me mantereí neutro neste debate taxonômico.

⁴⁰ Para uma introdução ao conceito de “função biológica”, cf. Garson, 2008.

o problema de sua origem evolutiva, do que propriamente quais são os animais conscientes e por quais métodos podemos identificar a sua presença.

Por mais que Andrews expresse que a abordagem biológica considera a consciência um fenômeno amplamente difundido nas mais variadas espécies, indo de mamíferos a nematoides, a sua apresentação desta abordagem não se encaixa exatamente como uma proposta de método de identificação da presença da consciência (ANDREWS, 2020a, p. 95). De fato, esse tipo de abordagem consistiria estaria mais próximo de ser um manifesto cujo objetivo é ampliar o leque de animais que devem ser estudados, do que particularmente uma solução para os problemas da distribuição e da mensuração. Nas palavras de Andrews:

Endereçando o problema da consciência pelo exame de sua função ao longo de seu tempo evolutivo, as abordagens biológicas sugerem que qualquer estudo fisiológico da consciência deve não limitar o seu foco aos primatas ou até mesmo aos mamíferos.⁴¹ (ANDREWS, 2020a, p. 95)

Isto posto, vejamos a seguir as três abordagens que propõem responder os problemas da distribuição e da mensuração da consciência animal.

1.3.1 A Abordagem da Percepção Direta

A abordagem da percepção direta⁴², também conhecida como abordagem não inferencialista, sustenta que a atribuição de consciência a um determinado animal é feita não por meio de inferências baseadas em seus comportamentos, mas em razão da percepção direta dos seus estados mentais conscientes. Da mesma forma que percebo diretamente a mesa diante de mim, de acordo com esta posição, perceberíamos a consciência nos animais. Este tipo de argumentação remonta a, pelo menos, o século XVIII com o filósofo David Hume (HUME, 2009). Segundo Hume, do mesmo modo que é evidente que os humanos são seres dotados de consciência, os animais também os são. Com efeito, bastaria prestarmos atenção para nos atentarmos ao fato dessa verdade evidente (HUME, 2009, pp. 209-212, livro 1, seção 16). No entanto, como meio de convencer o leitor, Hume também ofereceu argumentos que, de certo modo, se afastam da abordagem da percepção direta que ele mesmo endossa. Esses argumentos

⁴¹ “By addressing the question of consciousness by examining its function over evolutionary time, biological approaches suggest that any physiological study of consciousness should not limit its focus to primates or even to mammals.”

⁴² No que diz respeito à mente de outros humanos, Fred Dretske (1973) foi um dos pioneiros dessa abordagem. Mais recentemente, Shaun Gallagher (2008) e Roberto H. S. Pereira (2021) defenderam uma versão desse modelo.

se assemelham mais aos que Michael Tye (2017, 2018) e outros proponentes da abordagem newtoniana forneceram recentemente (FISCHER, 2021; THAGARD, 2021).⁴³ Nas palavras de Hume:

É com base na semelhança entre as ações externas dos animais e as por nós mesmos realizadas que julgamos que também suas ações internas se assemelham às nossas. E o mesmo princípio de raciocínio, levado um pouco adiante, nos fará concluir que, como nossas ações internas se assemelham umas às outras, as causas de que elas derivam também têm de ser semelhantes. Portanto, quando apresentamos uma hipótese para explicar uma operação mental comum aos homens e aos animais, devemos poder aplicar a mesma hipótese a ambos. (HUME, 2009, p. 210, § 3)

Creio que no parágrafo acima, Hume fornece dois argumentos analógicos. São eles:

- A) Os humanos *H* exibem um comportamento *C* e possuem estados internos *E*.
- B) O animal não humano *S* exibe um comportamento *C'* que é semelhante a *C* e possui estados internos *E'*.
- C) Ora, como o comportamento *C* exibido por *H* e *C'* exibido por *S* são semelhantes, os estados internos *E'* de *S* também devem se assemelhar aos estados internos *E* de *H*. Por analogia infere-se que:

-
- D) *S* possui estados internos *E'* que se assemelham aos estados internos *E* de *H*.

Assumindo que os estados internos a que Hume se refere sejam estados mentais conscientes, segue-se que:

- A) Os humanos *H* possuem estados internos conscientes *E* que causam o comportamento *C*.
- B) O animal não humano *S* possui estados internos conscientes *E'* que se assemelham aos estados *E* que *H* possui.

⁴³ Ao que me parece, a principal diferença entre o argumento fornecido por Hume e por Tye é a de que Hume utiliza um argumento analógico, ao passo que Tye apresenta um argumento abduutivo. Para mais informações sobre a posição de Tye e no que consiste o argumento analógico e abduutivo, cf. o capítulo 3 desta dissertação.

Ora, dado que os estados E' de S se assemelham aos estados E de H . Pode-se inferir por analogia que:

- C) Os estados E' de S causaram um comportamento C' que é semelhante ao comportamento C que H exhibe.

O 1º argumento expressa que, como os animais e os humanos exibem comportamentos semelhantes e ambos possuem estados internos, por analogia poder-se-ia inferir que os estados internos dos animais se assemelham aos dos humanos. O 2º argumento, por sua vez, expressa que os animais possuem estados internos conscientes que se assemelham aos que os humanos possuem. Esses mesmos estados internos conscientes dos humanos causam certos comportamentos. Ora, por analogia poder-se-ia inferir que certos comportamentos dos animais também são causados por estados internos conscientes, uma vez que entre os animais e humanos esses estados se assemelham. No 3º capítulo desta dissertação fornecerei maiores detalhes sobre os argumentos analógicos, como os empregados por Hume acima.

De acordo com Colin Allen (2004), Dale Jamieson (1998) e John Searle (2010) são dois dos filósofos que sustentaram mais recentemente posições similares a de Hume. Eles sustentam que, seja em virtude da nossa interação com os animais, ou por meio da percepção das suas expressões, de alguma forma percebemos diretamente a presença da consciência neles (JAMIESON, 1998; SEARLE, 2010). Segundo este modelo, quando percebemos uma expressão E_n , como, por exemplo, um grito de dor ou a expressão facial de surpresa de um chimpanzé, percebemos diretamente que este é um animal dotado de consciência. Ou seja, E_n não suscitaria uma sequência de premissas na nossa mente que garantiriam a inferência indutiva de que x possui consciência. De algum modo, por meio de E_n , perceberíamos diretamente, isto é, sem a utilização de uma inferência, que x tem consciência.⁴⁴

Por outro lado, devido a algum motivo obscuro, quando se está em um laboratório diante de uma placa de cultura bacteriana, não percebemos estes organismos como seres conscientes. Como bem Searle explicita, “[...] não importa realmente *como* eu sei que meu

⁴⁴ Alguns animais, como os cachorros (*Canis familiaris*), são capazes de inferir o estado emocional através da visualização da expressão facial exibida pelo humano e, assim, utilizar essas informações emocionais para tomar decisões. Por exemplo, foi observado que os cachorros tendem a preferir interagir e pedir ajuda a humanos que anteriormente apresentaram expressões faciais com valência positiva, como as de felicidade, do que aqueles que apresentaram com valência negativa, como as de raiva (ALBURQUERQUE et al, 2022). Claro, apesar desta capacidade denotar uma alta complexidade cognitiva do animal, ela não implica diretamente na tese de que os cachorros possuem algum mecanismo inato capaz de detectar estados mentais conscientes alheios, como certos proponentes da abordagem da percepção direta poderiam sustentar.

cachorro é consciente ou mesmo *se* eu “sei” ou não que ele é consciente. O fato é que ele é consciente, e a epistemologia nessa área deve *partir* desse fato.” (SEARLE, 2010, p. 117). De forma similar ao contador de Geiger, um instrumento capaz de detectar a radiação no ambiente, talvez possuamos uma ferramenta mental inata no cérebro que funciona como um detector de consciência, aquilo o que Andrews chamou de “detector animista” (ANDREWS, 2020a, p. 25). Ao entrar em contato com, por exemplo, um cachorro, o detector animista soaria o alarme em nossos cérebros de que o animal em questão possui estados mentais conscientes. Porém, o detector não faria o mesmo diante de uma bactéria ou ameba, uma vez que esses organismos não seriam conscientes. Todo esse processo se realizaria de forma imediata na mente humana, sem o intermédio de inferências.

No entanto, poderia ser posta uma diferença relevante entre o contador de Geiger e o detector animista que enfraqueceria a analogia entre os dois. O contador de Geiger fornece uma mensuração da radiação média no ambiente, isto é, ele não produz resultados binários do tipo 0 ou 1. Por exemplo, o contador de Geiger pode fornecer como valor um número como 16.22 $\mu\text{Sv/h}$. Ademais, por mais mínima que seja, sempre haverá um nível de radiação passível de ser detectável no ambiente.⁴⁵ Como nem sempre o detector animista detecta a consciência, vide o caso da placa de cultura bacteriana, então ele não funcionaria como um contador de Geiger. Contudo, o endosso de uma hipótese gradualista ou tudo-ou-nada sobre a natureza da consciência mudaria como o detector animista funciona. Por exemplo, caso se adote uma tese do tipo tudo-ou-nada, o detector animista não poderia admitir variações entre os estados conscientes diretamente percebidos, pois essa tese assume uma postura binária quanto à consciência. Segundo esse modelo, ou o detector animista detectaria a consciência ou não a detectaria, uma vez que a consciência seria um fenômeno que está ou não presente no animal em questão. Neste sentido, o detector animista só captaria 0 e 1, dado que a consciência seria um fenômeno binário. Por outro lado, o endosso de uma tese gradualista possibilitaria que o detector animista constatasse gradações de estados mentais conscientes. Ora, na medida em que um gradualista, seja o do tipo *scala naturae* ou multidimensional,⁴⁶ considera a consciência um fenômeno que admite gradações, o detector animista poderia, assim como o contador de Geiger realiza com a radiação, indicar que os animais possuem estados conscientes gradativos. O que se tem aqui é um detector não binário, onde para além do 0 e 1 existiria uma série, ordenada ou

⁴⁵ Evidentemente, a não detecção, pela razão que seja, de um nível mínimo de radiação por parte do contador de Geiger não anula o fato da radiação poder estar presente naquele ambiente. A ausência de evidência pode não ser evidência de ausência.

⁴⁶ Cf. o apêndice para visualizar as representações gráficas dos modelos gradualistas *scala naturae* e multidimensional.

não, de gradações de estados conscientes possivelmente existentes. Em suma, a depender da tese sobre a natureza da consciência endossada (tudo-ou-nada ou gradualismo), o detector animista também funcionaria de forma diferente (detector binário ou não binário).

Segundo a abordagem da percepção direta, não faria sentido exigirmos como forma de justificação argumentos a favor da percepção direta da consciência dos animais. Na medida em que é evidente que possuímos acesso imediato à consciência dos outros animais, o ônus da prova recai sobre aqueles que se manifestam céticos quanto à consciência animal. Dado que percebemos diretamente se o animal é consciente, caberia ao cético provar o contrário. A nossa percepção direta da consciência nos animais seria algo evidente (JAMIESON, 1998).

Todavia, não me parece claro que percebamos diretamente a consciência dos animais ou mesmo dos humanos. Considere o seguinte exemplo. A síndrome do encarceramento (*locked-in syndrome*) é um transtorno neurológico onde o sujeito perde completamente o controle dos seus músculos voluntários, deixando intacto apenas o controle da movimentação dos olhos. Ela ocorre quando há uma lesão na região do tronco encefálico conhecido como ponte. Com a lesão, é interrompido os caminhos entre os *outputs* do córtex com a medula espinhal. Os *outputs* vindos do córtex não conseguem chegar à medula espinhal, causando a paralização integral dos músculos do indivíduo. Por exemplo, devido à lesão na medula espinhal, o *output* de comando do braço enviado do córtex não chega ao seu destino, o que impossibilita o indivíduo de executar tal movimento. Apesar disso, a consciência do indivíduo permanece intacta. O indivíduo se encontra em um estado de aprisionamento dentro do seu corpo, consciente como uma pessoa neurotípica, mas sem a capacidade de movimentação (a não ser dos olhos) (DEHAENE, 2014; GAZZANIGA, 2018; LEÓN-CARRIÓN et al, 2005). Um caso famoso é o do jornalista francês Jean-Dominique Bauby, que após sofrer um acidente vascular cerebral (AVC) acorda no hospital com tal síndrome. O seu único movimento voluntário era a capacidade de piscar um de seus olhos. Foi com apenas este movimento que ele realizou a extraordinária tarefa de ditar um livro inteiro (DEHAENE, 2014).

Dado o exemplo supracitado, o nosso detector animista diria que devido apenas à percepção do sujeito deitado em uma cama, incapacitado de exercer quase todo tipo de comportamento, nos faltaria a evidência imediata de que ele é um indivíduo dotado de consciência. Neste caso, o detector falharia na sua função de captar imediatamente a consciência que se encontra presente no indivíduo. O proponente da abordagem da percepção direta poderia replicar que a evidência imediata da consciência estaria garantida por meio da movimentação dos olhos. Entretanto, até mesmo os pacientes em estado vegetativo possuem

essa e outras capacidades.⁴⁷ Claro, disso não se segue que o proponente da tese da percepção direta deveria então aceitar que os indivíduos em estado vegetativo que apresentam movimentos oculares espasmódicos são conscientes. Na verdade, esse e outros tipos de movimentos involuntários ocorrem geralmente por mero reflexo (DEHAENE, 2014). Porém, a dúvida ainda permaneceria em casos ainda mais radicais, como os de síndrome do encarceramento completo (*complete locked-in syndrome*). Os pacientes com essa versão ainda mais severa da síndrome possuem inclusive deficiências oculomotoras, o que os privam de até mesmo moverem os seus olhos. Isto acaba por resultar na incapacidade do estabelecimento de quase todo tipo de comunicação. Apesar disso, alguns autores sugeriram que seria possível nos comunicarmos com esses pacientes através de interfaces cérebro-computador, como os que utilizam o Eletroencefalograma (EEG), um método não invasivo que mede o potencial elétrico da atividade cerebral que foge do escalpo (GUGER et al, 2017). Em um estudo recente, pesquisadores conseguiram estabelecer a comunicação com um paciente portador de esclerose lateral amiotrófica (ELA) que se encontrava com a síndrome do encarceramento completo. A ELA, grosso modo, consiste em uma condição neurodegenerativa severa que aflige os neurônios motores fazendo com que o indivíduo perca o controle dos seus músculos voluntários. Por meio da inserção de dois microeletrodos no córtex motor suplementar e primário (Figura 4), os pesquisadores conseguiram se comunicar com o paciente conforme ele gerava tons de sons diferente através da modulação da ativação neuronal dessas áreas. Assim, conforme o paciente voluntariamente modulava, por exemplo, a ativação de um certo grupo neuronal do córtex motor suplementar, ele gerava um determinado tom sonoro. Como *output*, esse tom sonoro fazia com que letras fossem soletradas em uma interface cérebro-computador formando palavras e frases semanticamente corretas. Era através desse modo que o paciente completamente destituído de movimentos voluntários mantinha a comunicação com os pesquisadores (CHAUDHARY et al, 2022). Ora, se dependêssemos do nosso detector animista, Bauby provavelmente não ditaria o seu livro e o paciente portador de ELA nunca se comunicaria com os pesquisadores, pois eles seriam considerados pelo detector animista indivíduos destituídos de consciência.

Os defensores desta abordagem poderiam argumentar que este é um caso extremo de um transtorno neurológico raro entre os humanos e que não se refletiria nos animais. No entanto, mesmo entre os animais o nosso detector poderia falhar ou ser impreciso. Pode ser

⁴⁷ Existe um frutífero debate sobre se existem pacientes em estado vegetativo que apresentam estados conscientes, como no caso dos sonhos REM. Também há uma discussão sobre qual seria o melhor método capaz de determinar isso. Cf. Dehaene, 2014; Klein, 2017; Owen et al, 2006; Shea e Bayne, 2010.

claro que mamíferos como os primatas, golfinhos e cachorros possuem consciência (de WAAL, 2017, 2021). Poderíamos até mesmo conceder como evidentemente conscientes as diversas espécies de pássaros dotadas de uma capacidade cognitiva complexa. Os corvídeos, por exemplo, são famosos por possuírem uma cognição avançada, chegando a rivalizar com a dos grandes primatas e outros macacos (BALAKHONOV e ROSE, 2017; PIKA et al, 2020). Algumas evidências apontam, por exemplo, que os corvídeos conseguem inibir um comportamento imediato para receber uma recompensa maior no futuro, construir ferramentas compostas e planejar o futuro através da armazenagem de comida em um certo local a fim de saciar a fome que terá no dia seguinte (von BAYERN et al, 2018; DUFOUR et al, 2012; GÜNTÜRKÜN e BUGNYAR, 2016; RABY et al, 2007).

Porém, como deveríamos tratar dos casos filogeneticamente mais distantes de nós, como os dos artrópodes que possuem uma cognição sofisticada? Existem evidências de que certas espécies de abelhas, por exemplo, desempenham comportamentos cognitivamente complexos como: a habilidade de supervisionar os seus próprios processos cognitivos, a utilização de fezes como ferramentas para inibir possíveis invasores de sua colônia e a capacidade de navegar no ambiente através do uso de mapas de memória espacial (CHITTKA, 2022; LOUKOLA et al, 2017; MATILLA et al, 2020; MENZEL et al, 2005; PERRY e BARRON, 2013). Outras evidências sugerem que as abelhas poderiam realizar até mesmo estados afetivos como os de dor e ansiedade (BATESON et al, 2011; GIBBONS et al, 2022). Seriam elas conscientes segundo a abordagem da percepção direta?

Considere então os cefalópodes, como os polvo e os chocos. Estes são animais que possuem uma inteligência notável, embora possuam uma arquitetura cognitiva e cerebral totalmente diferente da nossa (GODFREY-SMITH, 2013, 2019a; SCHNELL et al, 2020). Os polvos, por exemplo, possuem uma quantidade maior de neurônios localizados em seus tentáculos do que especificamente em seu cérebro central (HOCHNER et al, 2006; ZULLO e HOCHNER, 2011). Apesar desta diferença, algumas evidências apontam que os polvos possuem um estado análogo ao do sono REM⁴⁸ e que provavelmente sentem dor (CROOK, 2021; MEDEIROS et al, 2021). De forma similar aos corvídeos, os chocos também conseguiriam inibir um comportamento imediato como meio de obter uma recompensa maior no futuro (SCHNELL et al, 2021).

⁴⁸ Existe evidência de que até mesmo a aranha-saltadora (*Evarcha arcuata*) possui um estado análogo ao do sono REM (RÖßLER et al, 2022).

Alguns autores defendem que até mesmo certas bactérias, como as *Escherichia Coli*, possuiriam alguma forma de proto-cognição (GODFREY-SMITH, 2016; LYON, 2015). O nosso detector animista teria como afirmar que também elas possuem consciência? Isto posto, muitos sustentariam que sim, todos esses organismos citados seriam claramente conscientes. Pense, por exemplo, em um biopsiquista ou um pampsiquista (ANDREWS, 2018; CHALMERS, 2017; GOFF, 2017; HAECKEL, 1892; REBER, 2016; THOMPSON, 2022). Creio não ficar claro o que a abordagem da percepção direta diria a respeito de todos esses casos.

Apesar de acreditar não ter oferecido argumentos cabais que refutam essa abordagem, considero razoável que, dentre todas as outras abordagens apresentadas, esta seja a menos plausível como resposta ao problema da mensuração. Na medida em que é endossado esse tipo de abordagem, o problema é trivializado. A consciência dos outros animais se torna algo óbvio e evidente, restando pouca margem para discussões. Deste modo, a desconsiderarei para nos próximos capítulos tratar mais a fundo as outras duas abordagens, a saber, a teórica e a epistêmica.

1.3.2 A Abordagem Teórica

A abordagem teórica propõe como forma de solução do problema da mensuração a aplicação tácita ou explícita de uma teoria prévia acerca da consciência humana aos animais. Segundo esse tipo de abordagem, devemos começar a investigação construindo ou escolhendo uma das inúmeras teorias da consciência humana disponíveis na literatura. É somente após a escolha da teoria que poderíamos verificar, mediante a utilização dela, se os outros animais possuem ou não consciência. Em resumo, o primeiro passo seria a formulação de uma teoria robusta da consciência nos seres humanos. Formulada então a teoria *T* da consciência em humanos, passasse agora a transportá-la diretamente para o caso dos animais.⁴⁹ Deste modo, seria por meio de *T*, que identificaríamos se um determinado animal ou espécie possuiria, ou não, consciência (ANDREWS, 2020a; ANDREWS e MONSÓ 2021; BIRCH, 2020a; SHEVLIN, 2021). Como Lau bem resume:

Uma boa maneira de resolver isso pode ser primeiro descobrir o que *em princípio* explica a consciência nos casos humanos. A partir daí, podemos observar se mecanismos semelhantes existem ou não nos animais e robôs. Isso

⁴⁹ Ou qualquer outro sistema cognitivo alvo. Como dito anteriormente, o foco desta dissertação é no reino animal (*Metazoa*).

não fornecerá nenhuma prova dedutiva, é claro. Ele assume que o nosso conhecimento da consciência em humanos está correto *e* que os princípios se generalizam para outras criaturas. Mas isso ainda é muito melhor do que palpites subjetivos. Para fazer essa generalização indutiva, **precisamos de uma teoria da consciência.**⁵⁰ (LAU, 2022, p. 151, negrito meu).

Note que, embora à ciência da consciência já seja uma área razoavelmente estabelecida, ainda há uma abundância de teorias diferentes e excludentes entre si sobre a consciência. Algumas pesquisas levantam que, somente dentre as neurobiológicas, existem cerca de 13 a 22 teorias relevantes da consciência (DOERIG et al, 2021; SETH e BAYNE, 2022). Isso denota uma falta de consenso científico que em outras áreas mais maduras, como a física e a biologia, dificilmente há - salvo quando se trata de assuntos que estão na fronteira do conhecimento, como a gravitação quântica e o nível, ou os níveis, em que a seleção natural opera - (GREENE, 2001; OKASHA, 2006). Ora, por que utilizar uma abordagem que já na sua origem é controversa, uma vez que ainda estamos distantes de um consenso científico mínimo sobre qual é a teoria da consciência correta? Penso que ao optarmos pela alternativa teórica estaríamos colocando a carroça na frente dos bois.

Outro seguinte questionamento poderia ser levantado: “como explicitar os mecanismos cognitivos identificados como constitutivos da consciência de acordo com a nossa teoria preferida, de modo a torná-los aplicáveis para além dos casos humanos.”⁵¹ (SHEVLIN, 2021, p. 4). Shevlin (2021) denominou essa questão de “o problema da especificidade” (*the specificity problem*). Trata-se de uma questão contida no problema da distribuição que diz respeito a como tornar as teorias da consciência humana viáveis para tratar de casos como os dos animais, da I. A. e até mesmo de bebês. Segundo uma versão radical da abordagem teórica, um animal possuiria consciência se e somente se ele cumprisse todos os requisitos de *T*. Por exemplo, para um animal *x* ser identificado como consciente, ele deverá possuir o mesmo tipo de arquitetura cognitiva ou configuração neuronal que *T* necessita em sua explicação da consciência. Se *T*, portanto, considerar metacognição, relato verbal e tomada de decisão como

⁵⁰ “A good way to address this may be to first figure out what **in principle** accounts for consciousness in the human cases. From there, we can see if similar mechanisms exist or not in the animals or robots. This will provide no deductive proof, of course. It assumes that our knowledge of consciousness in humans is correct, **and** that the principles generalize to other creatures. But this is still far better than subjective guesses. To make this inductive generalization, we need a theory of consciousness.” Sempre que houver trechos em negrito é porque o original está em itálico.

⁵¹ “[...] how to spell out the cognitive mechanisms identified as constitutive of consciousness according to our preferred theory in such a way as to make them applicable beyond the human case.”

aspectos constitutivos da consciência humana, serão estes os mecanismos que deverão estar presentes no animal. Por outro lado, em uma versão moderada da abordagem teórica, considera-se que se o animal em questão cumprir de maneira genérica os requisitos de *T*, ele poderá ser identificado como consciente. Por exemplo, se *T* exige a posse do compartilhamento global de informações para que o sistema possa ser considerado consciente, a versão moderada da abordagem teórica interpretará esta exigência de forma mais liberal para o caso dos animais. Shevlin (2021) identificou como conservadora (*conservatism*) e liberal (*liberalism*) as respectivas abordagens acima.

Shevlin (2021) apresenta duas versões adicionais da abordagem teórica. A primeira é a denominada “incrementalista” (*incrementalism*). Segundo ela, a consciência é um fenômeno que se manifesta em graus. De acordo com essa visão, devemos nos perguntar não se um animal possui ou não consciência, mas em que medida, ou grau, ele a possui (SHEVLIN, 2021). O incrementalismo está assentado nas teses gradualistas, onde a consciência é um fenômeno que se manifesta em diferentes níveis ou graus, seja motivado pela complexidade da sua arquitetura cognitiva e neuronal, seja pela quantidade de informações que o sistema consegue integrar, ou em razão de algum outro requisito explicitado por *T*. O incrementalismo pode estar baseado em uma teoria gradualista do tipo *scala naturae* ou multidimensional⁵² (BIRCH et al, 2020; GODFREY-SMITH, 2018b, 2020a; KOCH, 2020; TONONI e KOCH, 2015). Com efeito, devo deixar claro que a defesa de uma teoria gradualista da consciência não implicaria necessariamente o endosso de qualquer versão de pampsiquismo, biopsiquismo e zoopsiquismo (ANDREWS, 2018; CHALMERS, 2017; GODFREY-SMITH, 2016; GOFF, 2017; HAECKEL, 1892; THOMPSON, 2022).⁵³ Um determinado animal, segundo uma teoria gradualista não bio/zoo/pampsiquista, pode não possuir consciência alguma, caso não cumpra os requisitos de *T*.

⁵² Cf. nota 14 de rodapé.

⁵³ Grosso modo, o pampsiquismo sustenta que a consciência é uma propriedade intrínseca, fundamental e ubíqua da realidade. Nesse sentido, toda a matéria possuiria, em alguma medida, um caráter consciente. Por sua vez, o biopsiquismo defende a continuidade entre a vida e a consciência. Segundo essa posição, a consciência seria um fenômeno presente em todos e somente os seres vivos. Por fim, o zoopsiquismo sustenta que a consciência é um fenômeno presente entre todos e somente os animais (ANDREWS, 2018). Arthur Reber (2016, 2018a; REBER et al, 2022), por exemplo, sustenta que, com exceção das plantas, a consciência é um fenômeno presente entre todos os seres vivos, desde as bactérias a até os humanos. Reber (2018b) se mantém agnóstico quanto às plantas serem conscientes. É controverso se as plantas possuem consciência. Certos autores argumentam que algumas plantas possuem funções cognitivas análogas a aquelas presentes em animais, como, por exemplo, reconhecimento de parentes e comunicação. Segundo esses autores, uma vez que as plantas possuem mecanismos cognitivos análogos aos dos animais, soaria injustificável elas não serem consideradas conscientes, ao passo que normalmente os animais o são (CALVO, 2017; SEGUNDO-ORTIN e CALVO, 2022). Cf. também Godfrey-Smith, 2016; Goff et al, 2022; Thompson, 2022.

A última versão apresentada por Shevlin (2021) é a denominada “versão da rejeição” (*rejectionism*). Segundo esta abordagem, não existiria uma questão de fato (*fact of the matter*) no que diz respeito à aplicação de *T* em animais. As categorias que estão envolvidas em *T* não se aplicariam ou nada diriam a respeito dos sistemas cognitivos que estão para além dos casos dos humanos. Isso não quer dizer, por sua vez, que os animais não possuam consciência. Na verdade, existiria um tipo de indeterminação segundo a qual nunca poderíamos determinar se um animal, que não possua as mesmas categorias requisitadas por *T*, possuiria ou não consciência (CARRUTHERS, 2018a, 2018b, 2019, 2020a). No próximo capítulo, analisarei essa versão de forma mais aprofundada.

Muitos pesquisadores utilizam-se da abordagem teórica como o meio de identificar a presença da consciência no reino animal. Dentre as teorias mais utilizadas estão a AIR (acrônimo em inglês para *Attended Intermediate Representations*) (PRINZ, 2012, 2018), as teorias do espaço de trabalho global, como a teoria do espaço neuronal de trabalho global (BAARS, 1988; CARRUTHERS, 2018a, 2019; DEHAENE et al, 1998; DEHAENE, 2014, MASHOUR et al, 2020), as teorias de ordem superior (LAU, 2022) e a teoria do esquema de atenção (GRAZIANO e WEBB, 2017; GRAZIANO, 2019, 2022).⁵⁴

1.3.3 A Abordagem Epistêmica

Por fim, a abordagem epistêmica considera que não é necessário o comprometimento com *T* para respondermos o problema da mensuração e distribuição (ANDREWS, 2020a; ANDREWS e MONSÓ, 2021; BIRCH, 2020a; BROWNING e VEIT, 2020; SHEVLIN, 2021; TYE, 2017). Ela sustenta que, “não precisamos fazer grandes compromissos teóricos com antecedência sobre os mecanismos subjacentes da consciência.”⁵⁵ (SHEVLIN, 2021, p. 10). Segundo essa abordagem, o melhor método que possuímos para atribuir consciência aos animais é mediante fatores pré-teóricos, isto é, fatores que independem de um endosso preliminar de *T* (ANDREWS, 2020a; ANDREWS e MONSÓ, 2021). A abordagem epistêmica busca, deste modo, se esquivar de um comprometimento prévio com alguma teoria da consciência, seja ela de natureza especulativa, neurobiológica ou ambas. Não seria necessário endossar um princípio metafísico para atribuir consciência aos animais. Deve-se possuir um método que nos permita *justificar* a melhor hipótese acerca da presença ou não

⁵⁴ Birch (2020a) considera a tese de Klein e Barron uma abordagem teórica (BARRON e KLEIN, 2016; KLEIN e BARRON, 2016). Shevlin (2021), por sua vez, a considera uma abordagem epistêmica.

⁵⁵ “[...] *we need not make major theoretical commitments in advance about the underlying mechanisms of consciousness.*”

da consciência nos outros animais (TYE, 2017, p. 11). Ela funcionaria, portanto, como uma espécie de “princípio organizador” dos empreendimentos científicos-filosóficos (BROWNING e BIRCH, 2022). Isso não quer dizer, por sua vez, que os humanos não possuam importância nesse tipo de empreendimento e que, por consequência, poderíamos nos guiar utilizando apenas as nossas próprias intuições. A maior parte das propostas epistêmicas consideram que devemos iniciar a nossa pesquisa elencando os fatores que indicam, ou nos permitem identificar, a consciência em humanos. A partir da identificação desses aspectos em humanos, estabelece-se uma base sólida de fatores indicativos que se correlacionam com a consciência em humanos. Por fim, verifica-se a presença destes fatores nos animais.

Note que iniciar a investigação com os humanos não implica o compromisso com uma teoria da consciência humana. A razão pela qual devemos começar com os humanos se dá basicamente pelos seguintes motivos: o maior nível de certeza diante da existência da consciência em nossa própria espécie⁵⁶, a capacidade de realizarmos experimentos que forneçam evidências robustas da presença da consciência e a alta precisão dos relatos verbais e comportamentais que garantem as correlações entre comportamento, neurofisiologia e consciência. Este último motivo fornece um *benchmark* consistente para a investigação em outros sistemas cognitivos (EDELMAN e SETH, 2009).

Os fatores pré-teóricos, aqueles que nos ajudam a identificar a presença da consciência nos animais, variam de acordo com a versão da abordagem epistêmica adotada. No geral, elas concordam com a maior parte das evidências levantadas. Porém, em virtude dos seus próprios argumentos, as diferentes versões da abordagem epistêmica podem se diferenciar em alguns detalhes.

As propostas contidas epistêmicas se assentam, sobretudo, naquilo que chamarei de “*a estratégia dos marcadores de consciência*” (EMC). Em suma, a EMC propõe que a melhor forma de solucionarmos o problema da mensuração é utilizando os chamados “*marcadores de consciência*”. Dado o funcionamento das leis da natureza no mundo atual, os marcadores de consciência constituem um conjunto de indícios ou evidências que, por estarem relacionados à consciência, podem assinalar a presença da consciência no animal em questão (ANDREWS, 2020a; ANDREWS e MONSÓ, 2021; BAYNE e SHEA, 2020; SHEVLIN, 2021). Não devemos confundir os marcadores de consciência com os correlatos neurais da consciência (*neural correlates of consciousness*, NCC). O primeiro consiste em evidências da presença da

⁵⁶ Não entrarei no debate acerca do problema das outras mentes e em questões relacionadas às teses antirrealistas da consciência, como o ilusionismo e o eliminativismo. Para saber sobre tais posições, cf. Irvine e Sprevak (2020).

consciência em um dado animal, enquanto o segundo identifica uma determinada base neuronal como nomologicamente suficiente para a consciência (BIRCH, 2020a; CHALMERS, 2010a).⁵⁷ Ou seja, um NCC pode ser um marcador de consciência, mas o inverso pode não ser o caso.

Deve-se, portanto, identificar quais são os marcadores que se encontram ligados à presença da consciência nos humanos, para assim importá-los aos outros animais. Os marcadores de consciência podem ser comportamentais, neurofisiológicos ou até mesmo funcionais, como a realização de uma tarefa que dependa da consciência para ser bem executada. Como será visto no terceiro capítulo, cada versão da abordagem epistêmica utiliza de maneira distinta a EMC.

Isto posto, os alvos de discussão dos próximos capítulos serão respectivamente a abordagem teórica, especificamente a versão cético-radical de Peter Carruthers (2018a, 2018b, 2019, 2020a), e a abordagem epistêmica. Veremos que para Carruthers não existiria questão de fato acerca da presença e atribuição de consciência aos outros animais. As ciências e a filosofia da consciência deveriam, portanto, abandonar as pesquisas acerca do tema, pois elas nasceram fadadas ao fracasso. Com efeito, examinarei as motivações e os argumentos que Carruthers fornece para assegurar tal posição cético-radical sobre o tema. Buscarei mostrar que a sua posição possui contradições e que, ao contrário do que Carruthers prega, devemos continuar realizando pesquisas científicas e filosóficas sobre o assunto.

⁵⁷ Para saber mais sobre NCC, cf. Chalmers, 2010a; Crick e Koch, 1990; Morales e Lau, 2020.

2 **DEVEMOS NOS IMPORTAR COM A CONSCIÊNCIA ANIMAL!**

Atualmente, uma das posições que apresenta mais desafios à filosofia e à ciência da consciência animal é a proposta pelo filósofo e cientista cognitivo Peter Carruthers (2018a, 2018b, 2019, 2020a). Segundo o autor, *não* existiria questão de fato sobre se os animais possuem ou não estados mentais conscientes. Essa conclusão radical é produto da combinação entre duas teses por ele endossadas: a teoria do espaço de trabalho global e a estratégia do conceito fenomenal. Neste capítulo, avaliarei a posição cético-radical sustentada por Carruthers.⁵⁸ O capítulo está concebido da seguinte forma. Primeiro, apresentarei algumas noções propedêuticas fornecidas por Carruthers, como a concepção “tudo-ou-nada” da consciência e a distinção entre “realismo de *qualia*” e “irrealismo de *qualia*”. Em seguida, apresentarei brevemente no que consiste a teoria do espaço de trabalho global e a estratégia do conceito fenomenal. Posteriormente, reconstruirei os argumentos que embasam o seu ceticismo quanto à atribuição de consciência a animais. Por fim, fornecerei críticas que considero relevantes à posição de Carruthers. A primeira delas trata sobre um resultado implausível decorrente da sua conclusão radical, segundo a qual não haveria questão de fato para até mesmo organismos inteligentes capazes de solucionar o problema difícil da consciência, mas que possuem uma arquitetura cognitiva diferente da nossa. As outras duas críticas abordarão os compromissos controversos de Carruthers que geram contradições em sua posição. Diferentemente do que Carruthers (2020a, p. 18) prega, concluirei que devemos nos importar com a consciência animal.

2.1 **Introdução**

Como foi visto no capítulo anterior, no que se refere aos problemas da distribuição e da mensuração da consciência animal, existem ao menos 3 conjuntos de posições que buscam solucioná-las: a abordagem teórica, a epistêmica e a da percepção direta. A abordagem teórica sustenta que a melhor forma de determinar se um animal possui ou não consciência é mediante a aplicação de uma teoria da consciência humana. A abordagem epistêmica, por sua vez, apela a evidências relevantes que, por estarem de algum modo associadas à consciência, proveriam justificativas à hipótese *H* sobre se o animal *x* possui consciência. Por fim, a abordagem da

⁵⁸ Em sua taxonomia dos céticos quanto à consciência animal, Dung (2022) considera Carruthers (2019) um *agnóstico por princípio*. Essa posição afirma que *por princípio* não haveria como fornecer argumentos racionais e evidências empíricas a favor da hipótese *x* ou *y* quanto à atribuição de consciência aos animais. Embora considere correta tal caracterização, mantereí a minha nomenclatura “cético-radical”, pois creio que ela capta com maior precisão a posição de Carruthers.

percepção direta sustenta que percebemos diretamente os estados mentais conscientes através da interação com os seus comportamentos, tal como as expressões faciais e corporais.

Neste capítulo, avaliarei a posição de Carruthers que é uma das possíveis versões radicais da abordagem teórica. Segundo Carruthers (2019), a consequência da aplicação da melhor teoria da consciência humana ao caso dos animais é a ausência de questões de fato (*no fact of the matter*) sobre se eles possuem ou não consciência. Isso se dá em razão de uma *indeterminação de natureza semântica* dado o *descompasso* entre as categorias dispostas pela arquitetura cognitiva humana quando transportadas para as arquiteturas cognitivas dos animais (CARRUTHERS, 2019; SHEVLIN, 2021, p. 9). Não existiria, portanto, uma resposta “factual” às perguntas sobre a presença da consciência em animais, isto é, não haveria fatos para serem descobertos (CARRUTHERS, 2019, p. 164). Uma das consequências da posição de Carruthers é a irrelevância do próprio tema. Toda a atual explosão de pesquisas dedicadas à consciência animal consistiria, segundo o autor, em uma perda de tempo e recursos investidos em um programa de pesquisa equivocado (CARRUTHERS, 2018b, p. 58).⁵⁹ Tendo isso em vista, a sugestão de Carruthers (2019, p. 164) é o direcionamento das pesquisas científico-filosóficas para *apenas* o funcionamento dos processos cognitivos dos animais e o abandono das pesquisas sobre a consciência. A sua proposta pode ser resumida no seguinte lema retirada do título de um de seus artigos: “*Stop Caring About Consciousness*” (CARRUTHERS, 2020a).⁶⁰ Porém, como o título deste capítulo já alerta, argumentarei que, diferentemente do que Carruthers prega, devemos cientificamente e filosoficamente nos importar com a consciência animal.

No que toca a sua posição, Carruthers (2018a, 2018b, 2019, 2020a) a estabelece basicamente por meio da conjunção de duas teses. A primeira delas é a teoria do espaço de trabalho global (GWT) (BAARS, 1988; DEHAENE, 2014), considerada por muitos autores uma das principais teorias disponíveis da consciência (BLOCK, 2009; LAU, 2022; MYLOPOULOS, 2022; SETH e BAYNE, 2022). A segunda tese da conjunção é a “estratégia dos conceitos fenomenais” (ECF). Segundo Carruthers (2019, 2020a), a ECF consiste em uma

⁵⁹ Poder-se-ia argumentar que os experimentos realizados com animais, cujo objetivo consiste unicamente em investigar a cognição e a consciência, geralmente não são do tipo “*Rocket Science*”. Ou seja, não demandam uma grande quantidade de recursos financeiros para serem realizados. Assim, decretar os seus fins não se reverteria de fato em uma economia para as instituições de fomento (sobretudo as dos países desenvolvidos). Claro, permaneceria aberta a discussão sobre o aspecto ético envolvendo a manutenção de um programa de pesquisa com baixo custo, porém fundamentado em hipóteses supostamente equivocadas ou controversas.

⁶⁰ Interessante se notar que antes da sua virada ao representacionismo de primeira-ordem (CARRUTHERS, 2017b), Carruthers (2005) já concluía algo semelhante sobre a consciência animal. Os argumentos podem ser diferentes, mas a conclusão radical permanece a mesma.

maneira de responder aos diferentes argumentos antifisicistas. Isso se dá apelando a uma suposta classe peculiar de conceitos, os chamados “conceitos fenomenais (BALOG, 2012; STOLJAR, 2005).⁶¹ Seria através dessa conjunção que, segundo Carruthers, a GWT se tornaria uma explicação completamente reducionista da consciência, tal como o seu projeto naturalista objetiva. Como veremos, o resultado da conjunção será aquilo o que Jonathan Birch (2020b) denominou de “o condicional surpreendente de Carruthers”.

Dito isso, nas próximas seções fornecerei alguns prolegômenos importantes para o entendimento da posição de Carruthers. Apresentarei o que ele entende por “consciência” e a sua crucial distinção entre realismo de *qualia* e irrealismo de *qualia*. Em seguida, apresentarei sucintamente no que consiste a GWT e a versão particular da ECF por ele endossada. Logo após, reconstruirei os dois argumentos centrais que embasam a sua tese radical. Por fim, tecerei minhas críticas à posição dele.

2.2 A Consciência como um Fenômeno Tudo-ou-Nada

Antes de abordar propriamente o assunto desta seção, realizarei uma breve explanação com o objetivo de situar o leitor no debate. Para Carruthers (2018a, 2019), “consciência fenomenal” é um conceito caracterizado como sendo *fundamentalmente em primeira-pessoa*. Por mais que se tente utilizar locuções em terceira-pessoa, como: “há algo para alguém que é como estar embriagado”, ou exemplos do tipo “a experiência prazerosa de cheirar um café” e “a experiência desconfortável de calor em um dia do verão carioca”, a única função que eles exercem é a de chamar a atenção para a própria experiência fenomenal do indivíduo acerca daquele determinado estado. Ou seja, quando digo: “há algo para mim que é como sentir dor”, esta locução em terceira-pessoa intenta nada mais que chamar a atenção para o meu próprio estado afetivo em primeira-pessoa. De acordo com essa visão, o conceito de consciência se formaria baseado no contato (*acquaintance*) direto com uma certa experiência consciente. Em outras palavras, é o contato com a experiência consciente de um certo modo que proporciona ao indivíduo acesso ao entendimento com base em um determinado conceito fenomenal. É o contato (*acquaintance*) com a própria experiência consciente, portanto, que constitui a condição necessária ao entendimento de um conceito fenomenal C_n .⁶²

⁶¹ Note que mesmo os dualistas ou pampsiquistas podem admitir a existência de conceitos fenomenais, desde que fossem realizadas as devidas mudanças. Por exemplo, para o dualista de substâncias o referente do conceito fenomenal deve ser uma propriedade não física. Cf. Chalmers, 2010c, 2010d; Goff, 2017.

⁶² Na seção 2.4 apresentarei detalhadamente a concepção de Carruthers sobre o que é um conceito fenomenal.

Dito isso, Carruthers (2018a, 2019) considera a natureza da consciência como sendo um fenômeno do tipo tudo-ou-nada. Grosso modo, essa posição sustenta a seguinte disjunção: dado um sistema cognitivo S , ou S é consciente ou S não é consciente. Ou seja, S é consciente se e somente se ele instancia a consciência, caso o contrário S não é consciente. Segundo essa concepção, a consciência seria um fenômeno binário, onde ela está definitivamente presente (1) ou não a está (0). A consciência, portanto, não admitiria gradações e casos limítrofes (*borderline cases*) cuja presença de estados mentais conscientes não é claramente presente ou ausente, como certos gradualistas defendem (GODFREY-SMITH, 2020c; SCHWITZGEBEL, 2021; TYE, 2021).⁶³

Considere, por exemplo, a propriedade de ser careca. Existem pessoas que claramente *são* carecas e outras que claramente *não* são carecas. Contudo, entre ser e não ser careca existe uma série de gradações possíveis em que os indivíduos podem se situar. Ademais, na escala de ser careca existe um ponto onde o indivíduo pode cair em uma espécie de “limbo”. Neste limbo ele pode ser caracterizado como não sendo nem realmente careca e nem realmente cabeludo. Ele é *como se fosse* careca. Isto significa dizer que o conceito de careca é vago, pois ele admite casos limítrofes onde não se sabe exatamente se um sujeito possui ou não a propriedade em questão. Agora considere o conceito de “quadrado”. Ora, “quadrado” não seria um conceito vago, uma vez que não seria concebível uma figura que é *como se fosse* um quadrado. Uma determinada figura geométrica é quadrada ou não. Não existe uma escala gradativa de ser quadrado e, portanto, limbos (MICHEL, 2021, p. 620). Nesse sentido, Carruthers (2019) considera a consciência fenomenal uma propriedade parecida com a do quadrado. Ou seja, assim como uma figura geométrica G é ou não é um quadrado, um indivíduo S possui ou não consciência. Não existiriam gradações e casos limítrofes na consciência. Para Carruthers (2018a, 2019), seria até mesmo difícil de se conceber casos em que um indivíduo é parcialmente, ou mais ou menos, consciente.

No entanto, Carruthers (2019) admite que há gradações naquilo que seria o *conteúdo* de um estado mental consciente. Ou seja, é plenamente intuitivo que alguém esteja mais ou menos *consciente de* (*conscious of*) um determinado estímulo. Por exemplo, posso estar mais ou menos consciente da vermelhidão do livro a minha frente, ou mais ou menos consciente do som ao lado do meu quarto. Todavia, aquilo que aqui admite gradações não é a consciência em si, mas antes aquilo do que tenho consciência (consciência de criatura transitiva). Por mais pobre que seja o estímulo percebido, ainda assim se está ou não o experienciando. Ainda há

⁶³ Nem todo gradualista defende a existência ou concebibilidade de casos limítrofes de consciência, e. g. Birch, 2021.

algo aqui que é definitivamente como perceber o vermelho do livro a minha frente, ou escutar o som ao lado do meu quarto. Por mais que o conteúdo desse estado possa variar, ainda estou ou não o experienciando.

Por sua vez, a consciência também admitiria gradações quando o indivíduo está acordando ou adormecendo. Dizemos que a consciência aumenta o seu grau de riqueza quando se está acordando, ao passo que dizemos que o grau de consciência está se empobrecendo, ou apagando, quando se está adormecendo. No entanto, novamente não é a consciência fenomenal que aqui está admitindo gradações, mas aquilo o que denominam de “consciência de criatura intransitiva”. Isto é, o fato de estar acordado em oposição a estar dormindo ou em coma. Em ambos os casos supracitados o indivíduo estaria consciente no sentido fenomenal, pois definitivamente há algo para ele que é como estar acordando ou adormecendo. Em outros termos, o conteúdo de estar acordando ou adormecendo admite gradações, porém ainda há definitivamente algo para o indivíduo que é como estar acordando ou adormecendo. No último sentido, o da consciência fenomenal, não seria admitido gradações.

Portanto, “[...] não importa o quão empobrecido sejam seus conteúdos, definitivamente terá *algo como* estar neles”⁶⁴ (CARRUTHERS, 2019, p. 22). Desta forma, Carruthers conclui que a consciência fenomenal é um fenômeno tudo-ou-nada.

2.2.1 Realismo de *qualia* e Irrealismo de *qualia*

Em razão dos diversos argumentos antifisicistas, (e. g., a lacuna explanatória, o quarto da cientista Mary e a concebibilidade dos zumbis), alguns autores argumentam que a consciência fenomenal envolveria propriedades especiais, os chamados “*qualia*” (CHALMERS, 1996; JACKSON, 1982).⁶⁵ Carruthers (2018a, 2019) denominou a posição que defende a existência de tais propriedades especiais de “*realismo de qualia*”. Segundo essa posição, os *qualia* seriam propriedades intrínsecas e irredutíveis a qualquer estado físico, seja ele de natureza representacional, funcional ou neuronal. Ademais, os *qualia* seriam propriedades privadas do sujeito das experiências. Ou seja, a experiência peculiar de um certo estado consciente *E* se restringiria a somente o sujeito que estiver experienciando *E*.

Apesar de serem infáveis, essas propriedades especiais poderiam ser acessadas diretamente via introspecção pelo sujeito das experiências. Seriam os *qualia* que forneceriam o aspecto fenomenal que os estados mentais conscientes possuem. Por exemplo, quando nos

⁶⁴ “[...] *no matter how impoverished their contents, it will be determinately like something to be in them.*”

⁶⁵ Cf. o 1º capítulo para saber brevemente no que consiste tais argumentos.

ferimos geralmente experienciamos um certo estado consciente desagradável, a saber, *dor*. Para o realismo de *qualia*, seriam essas propriedades intrínsecas e irreduzíveis que forneceriam a característica qualitativa desagradável que um estado como a dor possui. Isso ocorreria na medida em que os *qualia* se ligassem ao estado consciente físico correspondente à dor. Por exemplo, o estímulo nocivo real da lesão em minha mão acarreta o disparo de uma certa população de neurônios no lobo parietal. O aspecto qualitativo da dor, por sua vez, não seria conferido apenas por meio dessa atividade neuronal específica, mas somente quando os *qualia* se ligassem a esse conjunto neuronal. Portanto, para que se tenha um estado qualitativo *Q* não basta que haja uma certa atividade neuronal *N*, pois o que confere a experiência subjetiva de *N* são os *qualia* ligados a *N*.

À primeira vista, o realismo de *qualia* poderia soar como uma espécie de dualismo implícito, onde os *qualia* seriam propriedades não físicas que garantiriam o aspecto qualitativo de certos estados físicos. Contudo, o compromisso com a existência dos *qualia* não implica imediatamente no dualismo. Segundo o próprio Carruthers (2019), o realismo de *qualia* é compatível com diversas formas de fisicismo, bem como as de dualismo.

Tendo isso em vista, Carruthers (2019, p. 7) afirma que o realismo de *qualia* se apresenta basicamente sob duas formas: o epifenomenalismo e as diversas espécies de monismo russelliano e pampsiquismo.⁶⁶ Segundo o epifenomenalismo, os *qualia* seriam propriedades epifenomenais, isto é, não teriam poderes causais sob o mundo físico. Em outros termos, os *qualia* seriam propriedades supervenientes a estados físicos, mas causalmente irrelevantes, uma vez que não exercem nenhum impacto em nosso mundo. Diversos problemas emergem do epifenomenalismo. Porém, o que aparenta colocar mais dificuldades a ele é o chamado “fecho causal do mundo físico” (*physical causal closure*). Segundo o fecho causal do mundo físico, para todo evento físico existe uma e apenas uma causa física (salvo casos raros de sobredeterminação genuína). Entretanto, se os *qualia* forem propriedades não físicas e causalmente inertes, eles não poderiam ser capazes de causar eventos no mundo físico (PAPINEAU, 2002). Não seriam os *qualia* que causariam, por exemplo, os nossos estados mentais tais como crenças e desejos - mesmo assumindo que eles sejam eventos físicos realizados em nosso cérebro, como algumas evidências sugerem – (CONNORS e HALLIGAN,

⁶⁶ Acredito que as posições denominadas por Pautz (2021, cap. 2) de “visão do estado físico interno” também sejam realistas de *qualia*. A tese da visão qualitativa de Papineau (2021), por exemplo, seria uma delas. Em linhas gerais, a visão qualitativa sustenta que as propriedades conscientes são propriedades intrínsecas das pessoas e contingentemente relacionais com eventos externos, o que vai de encontro ao representacionismo (PAPINEAU, 2021, p. 83). Os estados mentais conscientes seriam, portanto, determinados pelas propriedades internas e intrínsecas dos indivíduos (como certos padrões neuronais), e não eventos externos (PAPINEAU, 2021, p. 50). Sobre críticas à posição de Papineau, cf. Pautz, no prelo; Pereira e Souza Filho, 2022.

2015; SEITZ e ANGEL, 2020). De acordo com Carruthers (2019, p. 8), mesmo as nossas crenças de que existe uma lacuna explanatória ou um problema difícil da consciência não seriam produtos dos *qualia*.

O monismo russelliano⁶⁷, por sua vez, sustenta que os *qualia* são quididades fundamentais do nosso mundo físico. Em linhas gerais, para essa posição os *qualia* poderiam ser entendidos como quididades intrínsecas e fundamentais da realidade que se auto-individuam. Em geral, os autores que advogam algum tipo de monismo russelliano ou pampsiquismo se comprometem, em algum grau, com a tese chamada “realismo estrutural epistêmico”.⁶⁸ Segundo o que eles argumentam, as ciências, como a física e a química, seriam incapazes de, *por princípio*, captar essas quididades. O motivo seria o de que mesmo as ciências mais fundamentais, como a mecânica quântica, só dizem respeito ao nível das estruturas e relações que compõem a realidade, deixando assim escapar as quididades. Colocando de outra maneira, as ciências naturais só dizem respeito à natureza das propriedades extrínsecas, disposicionais e/ou relacionais da matéria (GOFF, 2017, pp. 142-147). As quididades seriam aquilo o que fundamentariam as propriedades relacionais e estruturais da realidade. Nesse sentido, seriam as quididades que constituiriam a consciência, ao menos na versão pampsiquista do monismo russelliano (CHALMERS, 2017; GOFF, 2017). Para Carruthers, essas posições seriam realistas de *qualia*, pois elas consideram os *qualia* como sendo quididades extras e irreduzíveis a qualquer estado físico. Entre as razões pela qual Carruthers (2019, p. 9) rejeita tais posições, está o fato de que elas não acrescentam nada de novo à física fundamental, como possíveis predições empíricas. Ou seja, essas posições não desempenhariam nenhum papel explanatório relevante para as ciências naturais e, portanto, devem ser abandonadas.

Assim, Carruthers (2018a, 2019) rejeita o realismo de *qualia* e se compromete com o denominado “irrealismo de *qualia*”. O irrealismo de *qualia* pode também se apresentar de diversas maneiras, no entanto o que as diferentes versões possuem em comum é a busca em identificar a consciência com alguma forma de propriedade natural física. Em contraste com a posição realista, o irrealista de *qualia* endossaria algum tipo de reducionismo, isto é, ele visaria reduzir a consciência a propriedades físicas e/ou fisicamente realizadas. É nesse sentido que

⁶⁷ Chalmers (2017) realiza uma taxonomia detalhada entre os diversos tipos de posições russellianas, como: o pampsiquismo russelliano, o pamprotopsiquismo russelliano, o monismo russelliano e as versões constitutivas e não constitutivas de cada uma delas. Não versarei sobre cada uma dessas variações, pois fugiria do escopo desta dissertação. Cf. Chalmers, 2017; Goff, 2017; Goff et al, 2022.

⁶⁸ Também existe o realismo estrutural ôntico. Grosso modo, essa posição sustenta que aquilo o que existiria no nível mais fundamental da realidade seria tão somente estruturas e relações. A realidade mais fundamental, portanto, não seria composta por propriedades intrínsecas como os monistas russellianos e os pampsiquistas defendem. Cf. Ladyman e Ross, 2007.

Carruthers defende que a consciência fenomenal seria nada mais do que representações com conteúdo não conceitual transmitidas globalmente. Não existiriam propriedades especiais, extras, inefáveis, intrínsecas à própria experiência e *irreduzíveis* a propriedades e processos físicos. Em outras palavras, para Carruthers, a consciência seria nada mais do que representações que adentram o espaço de trabalho e ali são globalmente compartilhadas aos mais diversos tipos de sistemas consumidores, como, por exemplo, a ação racional, o relato verbal e a metacognição. Com efeito, a consciência poderia ser completamente explicada e reduzida através de processos informacionais como as que a teoria do espaço de trabalho global propõe (CARRUTHERS, 2019, p. 13). Embora Carruthers (2019, pp. 14-15) considere o irrealismo de *qualia* próximo ao ilusionismo, ele rejeita tal rótulo. O ilusionismo, em linhas gerais, é uma tese que nega a existência de propriedades fenomenais, argumentando que a aparência qualitativa que a experiência consciente possui é uma ilusão provocada pela introspecção (FRANKISH, 2016). O motivo para o rejeite, segundo Carruthers, é que ilusões são geradas espontaneamente, enquanto a crença de que existem *qualia* depende da articulação explícita de certos modos de reflexão e da construção de argumentos filosóficos.

Ainda assim o irrealismo de *qualia* de Carruthers estaria sujeito aos desafios impostos pelos argumentos antifisicistas, visto que poderíamos, por exemplo, conceber zumbis cujas representações não conceituais adentram o espaço de trabalho global, mas que não são acompanhadas por quaisquer experiências qualitativas. A resposta de Carruthers a esses desafios é a adoção da chamada “estratégia do conceito fenomenal”.⁶⁹ Segundo Carruthers (2019, 2020a), a conjunção entre a teoria do espaço de trabalho global e a estratégia do conceito fenomenal forneceria a explicação correta e completamente reducionista da consciência, ao menos nos casos de humanos. Dito isso, vejamos a seguir no que consiste a teoria do espaço de trabalho global.

2.3 A Teoria do Espaço de Trabalho Global

Para muitos autores, a teoria do espaço de trabalho global (GWT) figura dentre as principais teorias que atualmente versam sobre a consciência humana (BLOCK, 2009; GAZZANIGA, 2018; MYLOPOULOS, 2022; SETH e BAYNE, 2022).⁷⁰ A GWT foi, ao

⁶⁹ De acordo com a taxonomia de Chalmers (2010b), a posição de Carruthers (2019) poderia ser caracterizada como um materialismo tipo-B quanto aos argumentos antifisicistas. Grosso modo, o materialismo tipo-B é caracterizado pela combinação de um monismo ontológico com um dualismo conceitual. Segundo essa posição, não existiriam conexões *a priori* entre a consciência fenomenal e os conceitos em terceira-pessoa.

⁷⁰ O intuito aqui é prover apenas uma breve, porém útil explanação sobre a GWT. Para informações mais extensas e detalhadas, cf. Dehaene, 2014; Mashour et al, 2020. Sobre críticas à GWT, cf. Block, 2009; Lau, 2022.

menos da forma que se discute hoje, inicialmente desenvolvida nos anos 80 pelo célebre psicólogo cognitivo Bernard Baars (1988, 2005a). Nesta versão inicial, a GWT foi fundamentalmente uma teoria construída a partir de termos e conceitos cognitivos e psicológicos. Contudo, no final da década de 90 a GWT ganhou uma nova formulação sob a égide dos avanços nas neurociências. Esses avanços se deram, sobretudo, graças às pesquisas empíricas e às sofisticações teóricas dos neurocientistas Stanislas Dehaene, Michel Kerszberg e Jean-Pierre Changeux (1998). A esta nova versão deu-se o nome de “teoria do espaço neuronal de trabalho global” (DEHAENE et al, 1998; DEHAENE et al, 2001; DEHAENE e CHANGEUX, 2011; DEHAENE et al, 2017; DEHAENE, 2014; MASHOUR et al, 2020). Atualmente, é essa a versão que geralmente se debate. De forma independente, outros autores desenvolveram teorias semelhantes à GWT, como, por exemplo, Michael Tye (1995), que a concebeu nos termos da sua teoria conhecida pelo acrônimo “*PANIC*” (*poised, abstract, non-conceptual, intentional content*).⁷¹ Segundo Tye (1995, pp. 137-144), para que um conteúdo intencional seja consciente, ele deve satisfazer alguns critérios. São eles: a) estar apropriadamente disponível ao sistema cognitivo e desempenhar um papel causal nas crenças e desejos, b) ser intrinsecamente não conceitual e c) ser abstrato, ou seja, nenhum objeto particular concreto deve entrar em seus conteúdos.

Antes de prosseguir, devo esclarecer alguns termos. De acordo com a interpretação de Carruthers (2019, p. 97), quando se expressa que uma representação é *globalmente compartilhada*, isto *não* quer dizer que ela obrigatoriamente deve ser compartilhada a todos os sistemas consumidores. O autor entende que a representação deve ser compartilhada para muitos sistemas, sobretudo os que compõe as funções executivas (e. g. raciocínio, tomada de decisão e autocontrole). Em geral, os proponentes da GWT são vagos quando se trata de especificar quais são os sistemas metafisicamente relevantes e necessários para que a representação se torne consciente. Shevlin (2021) nomeia essa questão de “o problema da especificidade”. Quanto ao que exatamente seria o espaço de trabalho, Carruthers (2019, p. 97) o identifica à memória de trabalho. A memória de trabalho, grosso modo, seria um mecanismo mental responsável por armazenar e manipular temporariamente um conjunto de informações de modo a torná-lo disponível ao comportamento flexível do organismo. Ela é frequentemente associada a áreas mais frontais do cérebro, como o córtex pré-frontal dorsolateral (CURTIS e D’ESPOSITO, 2003; FUNASHI, 2017). Outras evidências sugerem que a memória de trabalho

⁷¹ Em seu livro de 1995 Tye nomeia a sua teoria de PANIC. Contudo, em seu livro mais recente, Tye (2021) passa a também adotar o nome “*Global Workspace*”.

também envolve atividades no córtex medial frontal e no lobo temporal medial, sendo neste último através de populações de “células conceito” (*concept cells*) que disparam em resposta a conceitos semânticos abstratos, como o de “pessoa” e “animal” (BORAN et al, 2022; KAMIŃSKI et al, 2017; KORNBLITH et al, 2017).

Isto posto, a GWT sustenta que um estado mental é consciente quando as representações adentram o espaço de trabalho global e ali são amplamente compartilhados entre os mais diversos sistemas consumidores da mente (Figura 9). Ou seja, quando as informações entram no espaço de trabalho global e são compartilhadas a sistemas como os de raciocínio, planejamento, tomada de decisão, memória, ação guiada, entre outros, elas se tornam estados conscientes (CARRUTHERS, 2019). Com efeito, a consciência seria “[...] apenas informação amplamente compartilhada pelo cérebro”⁷² (DEHAENE, 2014, p. 165, ênfase minha). Os estados mentais *inconscientes*, por sua vez, são aqueles que *não* adentram o espaço de trabalho global e, portanto, *não* são amplamente compartilhados no cérebro, embora possam exercer alguma influência no comportamento, vide os estados mentais subliminares de *priming* (DEHAENE, 2014).

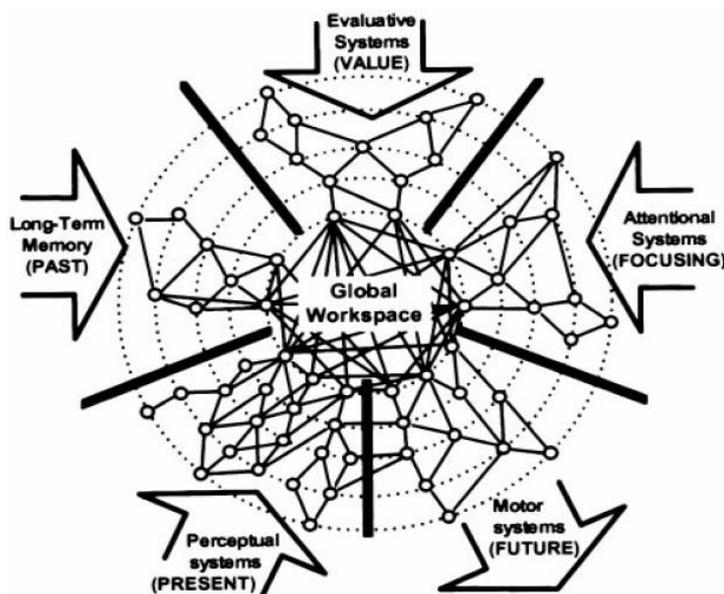


Figura 9 – Ilustração do espaço de trabalho global. Dehaene et al, 1998, p. 14530.

Em termos neuroanatômicos gerais, considera-se que os estados conscientes envolvem informações sendo transmitidas por neurônios com axônios que cobrem longas distâncias. Neste caso, seriam os neurônios do tipo piramidal contidos no córtex que contribuiriam de forma mais adequada no desempenho das tarefas responsáveis pela

⁷² “[...] just brain-wide information sharing.”

transmissão dessa informação. Apesar da maior densidade desses tipos de neurônios se encontrar no córtex pré-frontal, a GWT não a considera a base neuronal da consciência, pois isso a tornaria uma teoria localista, postura essa que os seus proponentes rejeitam (MASHOUR et al, 2020, p. 777). De fato, eles consideram que o córtex pré-frontal, em conjunto com o córtex parietal, é uma das áreas que desempenham um papel fundamental para a realização da consciência, uma vez que lá se encontra uma grande densidade de neurônios piramidais. Lembre-se que um estado mental E_n só se torna consciente quando é amplamente compartilhado no cérebro. Ou seja, E_n se torna consciente através do seu compartilhamento às mais diversas regiões cerebrais, indo do lobo frontal ao lobo parietal e temporal (DEHAENE, 2014; MASHOUR et al, 2020) (Figura 10).

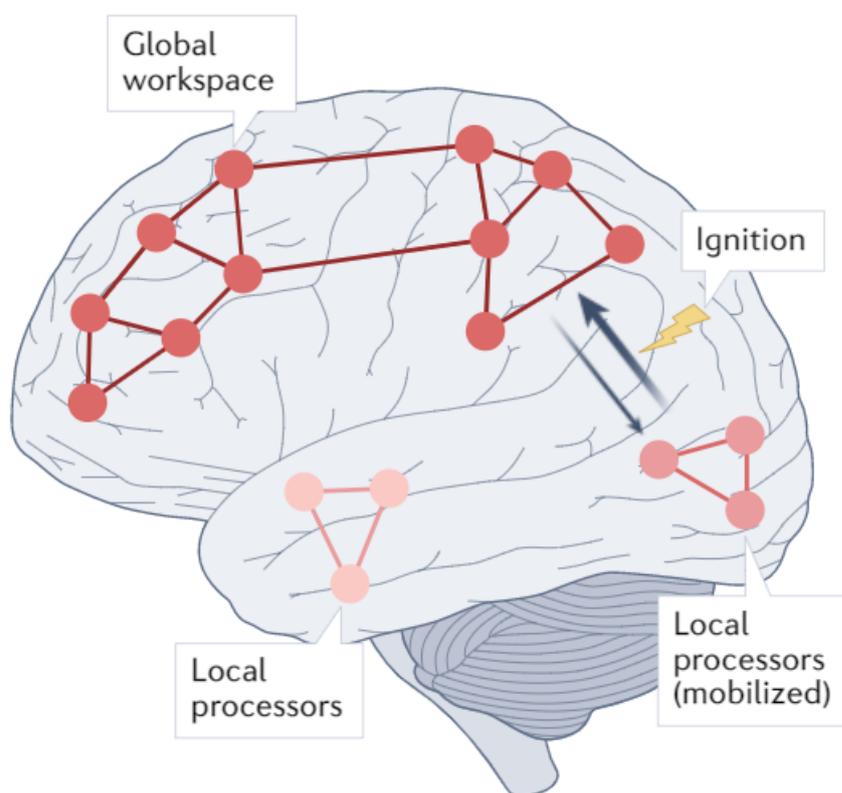


Figura 10 – Representação da GWT. Da esquerda para a direita: Espaço de trabalho global; Ignição; Processadores locais (mobilizados); Processadores Locais. Seth e Bayne, 2022, p. 6.

Ao menos na interpretação da GWT que aqui nos interessa, a consciência fenomenal seria nada mais que a própria consciência de acesso. Para Carruthers (2019), mais especificamente, a GWT identificaria a consciência fenomenal com um conteúdo não conceitual que ao adentrar o espaço de trabalho global é amplamente difundido entre os diversos sistemas dispostos em sua arquitetura cognitiva. Não haveria, portanto, dois fenômenos

empiricamente distintos, isto é, de um lado a consciência fenomenal e de outro a consciência de acesso. Como pode-se notar, Carruthers não reduz ou identifica a consciência diretamente a certos estados ou padrões neuronais, como Papineau (2021, pp. 95-97), por exemplo, o faz.⁷³ Embora também utilize as evidências neurocientíficas, Carruthers descreve a GWT em termos cognitivos e representacionais. Segundo ele, os dois motivos pelos quais não se deve explicar estados conscientes através apenas de mecanismos e processos cerebrais são os seguintes. Primeiro, existiria uma série de explicações de níveis distintos entre os fatos de ordem cerebral e os de ordem da consciência. Por exemplo, existem explicações de nível funcional, computacional e representacional. Segundo, as pessoas intuitivamente pensam que há uma distinção relevante entre os estados mentais e os estados físicos que resultam no tradicional problema mente-corpo (CARRUTHERS, 2019, pp. 117-118). O argumento pode ser resumido da seguinte maneira, se há diferentes níveis de explicações possíveis entre os fatos cerebrais e os fatos da consciência, e se as pessoas possuem intuições fortes que separam os estados mentais do físico, logo não haveria a necessidade, ao menos imediata, de se explicar a consciência em termos puramente cerebrais.

Segundo Carruthers (2019), a melhor forma de se explicar a consciência e se esquivar das fortes intuições dualistas das pessoas seria empregando termos mentais. Mais do que isso, a consciência poderia ser reduzida a conteúdos intencionais que desempenham um certo papel causal. Ou seja, a redução da consciência se daria na sua completa explicação em termos puramente de conteúdo intencional. Se seguíssemos as intuições naturalistas, o próximo passo seria a redução do conteúdo intencional a estados físicos de nível inferior, como certos estados ou padrões neuronais. Contudo, não é isso o que Carruthers propõe. Para Carruthers, representações e conteúdos representacionais são propriedades psicológicas reais que já foram plenamente naturalizadas através dos trabalhos das ciências da mente, sobretudo das ciências cognitivas e da psicologia. Deste modo, assim como Tyler Burge (2010) defendeu, Carruthers (2019, pp. 116-120) assume uma posição primitivista quanto a representações e conteúdos representacionais. As representações e os conteúdos representacionais seriam, segundo Carruthers, entidades amplamente utilizadas pelas ciências cognitivas, psicologia e até mesmo pela neurociência (WARD, 2020; DEHAENE, 2014), que desempenham um papel explanatório fundamental em suas pesquisas empíricas e teóricas. Por conseguinte, as representações e os conteúdos representacionais devem ser tratadas como entidades teoreticamente primitivas não

⁷³ Apesar de ser um fisicista, Papineau (2021, p. 90) considera a visão qualitativa compatível com o dualismo.

havendo, assim, necessidade de reduzi-las a propriedades físicas de nível inferior (como certos estados ou padrões neuronais).⁷⁴

Segundo essa interpretação da GWT, não existiria consciência fora do espaço de trabalho global. A GWT estaria, portanto, tacitamente em oposição ao chamado “argumento do transbordamento” (BLOCK, 2007a, 2011a).⁷⁵ Note que essa seria a interpretação irrealista de *qualia* da GWT, pois a consciência seria reduzível ao conteúdo não conceitual compartilhado globalmente, isto é, à consciência de acesso. Desta forma, não existiriam propriedades especiais e irreduzíveis, tais como os *qualia*, que ao se combinarem às representações disponíveis no espaço de trabalho global às confeririam o status de consciente no sentido fenomenal. Essa última interpretação seria realista de *qualia*, na medida em que admite a existência de tais propriedades. Segundo essa concepção, para que as representações se tornassem conscientes não bastaria apenas que elas adentrassem e estivessem disponíveis no espaço de trabalho global. As representações também deveriam se ligar aos *qualia*, uma vez que seriam estas últimas propriedades que proporcionariam o aspecto fenomenal às representações (CARRUTHERS, 2018a, pp. 197-198). Assim, para o realismo de *qualia* somente a conjunção entre representações globalmente compartilhadas e os *qualia* garantiriam o aspecto fenomenal experienciado pelo sujeito portador de tal estado.⁷⁶

No que tange as evidências que corroboram a GWT, geralmente elas são obtidas por meio dos experimentos do tipo contraste mínimo somado às técnicas de imageamento cerebral.⁷⁷ Grosso modo, a metodologia do contraste mínimo consiste na distinção entre estados conscientes e inconscientes através da realização de uma determinada técnica. Em geral é apresentado ao participante um certo estímulo sonoro, visual, tátil ou olfativo, de uma certa maneira que a torne consciente ou inconsciente. Ao mesmo tempo que isso ocorre, os pesquisadores, por meio de técnicas de imageamento, mensuram a atividade cerebral do participante. Um exemplo são os experimentos que utilizam o mascaramento. Neste tipo de experimento, um estímulo alvo é apresentado de forma breve entre dois estímulos “mascaradores”, sendo um que antecede e outro que sucede o estímulo alvo, de modo a torná-lo invisível ao participante. Por mais que o pesquisador diga ao participante que uma figura, digamos um rosto humano com a expressão de raiva, foi projetado entre duas claras imagens

⁷⁴ Cf. Burge (2010, cap. 8, 2022) para obter uma compreensão completa da posição primitivista sobre representação mental.

⁷⁵ Sobre o argumento do transbordamento, cf. o 1º capítulo desta dissertação.

⁷⁶ Tye (2021) me parece sustentar uma interpretação realista de *qualia* da GWT.

⁷⁷ Como dito anteriormente, não é o meu intuito aqui exaurir a literatura sobre o assunto. Para revisões a respeito das evidências que supostamente corroboram a GWT, cf. Carruthers, 2019, cap. 5; Dehaene, 2014; Mashour et al, 2020.

(que funcionam como mascaradores), ainda assim o participante não consegue perceber conscientemente (DEHAENE et al, 2001). Existiria, portanto, um certo limiar no qual a informação cruza e se torna consciente. Os resultados obtidos desses experimentos, normalmente via Eletroencefalograma (EEG), Ressonância Funcional Magnética (fMRI) e Magnetoencefalografia (MEG),⁷⁸ indicam que quando o estímulo é apresentado de modo consciente há uma intensa atividade distribuída nas diversas regiões do cérebro. Isso incluiria a ignição de atividades cerebrais desde o córtex V1 chegando a áreas como o córtex pré-frontal bilateral e regiões do lobo parietal. Dehaene (2014) denomina esse efeito cascata de uma das “assinaturas da consciência”. Os estímulos que foram tornados inconscientes ao sujeito através do mascaramento, por outro lado, são marcados geralmente por baixa atividade cerebral distribuída ao longo do córtex (Figura 11) (DEHAENE, 2014; MASHOUR et al, 2020).

Note que apenas as evidências supracitadas não garantem à GWT o status de teoria completamente reducionista da consciência, como Carruthers (2019, 2020a) afirma. Poderia ser o caso de a GWT ser apenas uma teoria de como são diferentemente realizados os estados mentais conscientes e inconscientes. Neste contexto, a GWT não consistiria em uma teoria completamente reducionista, mas apenas uma teoria descritiva da consciência. Ou seja, uma teoria dos correlatos neurais que a consciência possui. Essa seria, por exemplo, a posição de Tye (2021, pp. 87-88). Assim, as objeções antifisicistas, como a concebibilidade dos zumbis, a lacuna explanatória e o quarto de Mary (CHALMERS, 1996; JACKSON, 1986; LEVINE,

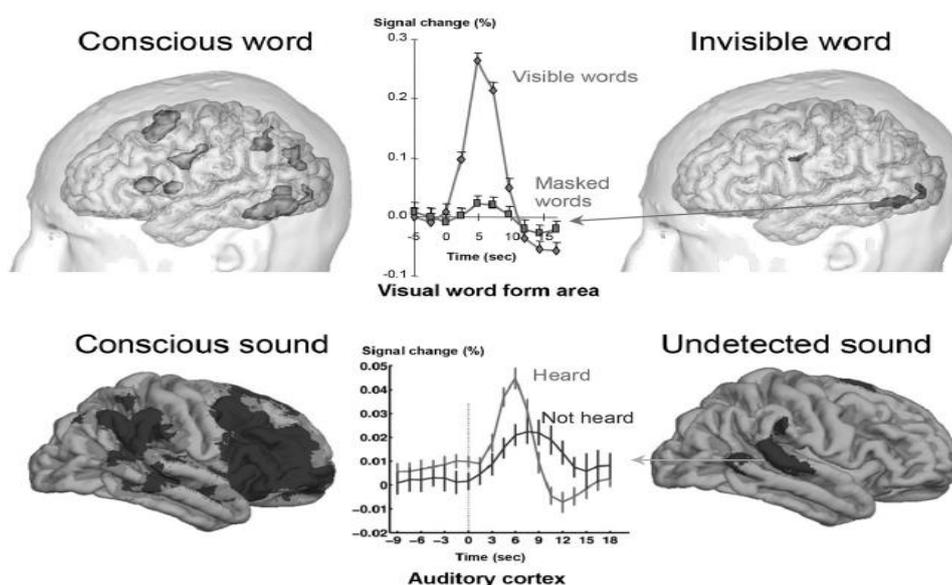


Figura 11 – Comparação realizada entre a atividade cerebral da percepção consciente e inconsciente de dois tipos de estímulos (palavras e sons). Dehaene, 2014, p. 119.

⁷⁸ Grosso modo, o EEG é um método não invasivo que mede o potencial elétrico da atividade cerebral. O fMRI, por sua vez, realiza a medição através dos níveis de oxigenação do sangue nas diferentes áreas do cérebro. Por fim, o MEG é um método que mede a atividade cerebral por meio dos sinais magnéticos gerados pela corrente elétrica que escapa do escalpo (BACKER, 2022).

1983), ainda permaneceriam de pé, uma vez que a GWT sozinha não conseguiria (ou mesmo intenta) respondê-las.

Contudo, como Carruthers afirma que a GWT é a teoria correta e completamente reducionista da consciência, ele deveria ser capaz de fornecer argumentos que justifiquem tal afirmação. Como forma de responder a esse questionamento, Carruthers se vale da chamada “estratégia do conceito fenomenal” (ECF). Segundo ele, a ECF dissolveria os argumentos antifisicistas e garantiria a redução da consciência a conteúdo não conceitual globalmente compartilhado. Seria a combinação entre a GWT e a ECF que, portanto, fundamentaria o seu projeto reducionista e naturalista da consciência (CARRUTHERS, 2018a, 2019, 2020a).

2.4 A Estratégia do Conceito Fenomenal

Como visto anteriormente, Carruthers defende que a GWT explica a consciência de forma completamente reducionista e naturalista. Se de fato a GWT é isso o que o autor propõe, ela não deveria ser apenas uma mera descrição correlativa entre a consciência e o compartilhamento global de informações, ou de um estado consciente e o seu correlato neural. Na verdade, a GWT deveria ser capaz de explicar todos os fatos relativos à consciência em termos psicológicos de terceira-pessoa (CARRUTHERS, 2020a). No entanto, segundo Carruthers (2019), para que o empreendimento reducionista obtenha sucesso ele deve responder com sucesso os argumentos antifisicistas (CHALMERS, 1996; JACKSON, 1986; LEVINE, 1983). Carruthers (2018a, 2019, 2020a) argumenta que é através da chamada “estratégia do conceito fenomenal” (ECF) que os argumentos antifisicistas se dissolveriam, garantindo assim o sucesso de seu empreendimento.

Embora a ECF não seja unanimidade na literatura fisicista (BALL, 2009; PEREIRA, 2019, 2020; TYE, 2009) e tampouco convença autores com posições heterodoxas, como os dualistas e pampsiquistas (CHALMERS, 2010b, 2010c; STOLJAR, 2005), ela provavelmente permanece sendo uma das estratégias mais populares entre os fisicistas, especialmente os de tipo-B (CHALMERS, 2010b; BALOG, 2012; CARRUTHERS, 2019, 2020a; PAPINEAU, 2002, 2020). Resumindo a proposta da ECF, Carruthers afirma que os argumentos antifisicistas surgem:

[...] não porque os nossos estados conscientes possuem propriedades especiais de qualquer tipo (*qualia*), mas em razão da maneira distinta em primeira-pessoa pela qual é possível *pensarmos acerca* desses estados. É porque esses pensamentos carecem de conexões conceituais com os físicos, funcionais ou

mesmo representacionais que podemos conceber a possibilidade de zumbis; e é também por isso que podemos ver que mesmo uma descrição física/funcional/intencional completa de nós mesmos não implicaria a verdade desses pensamentos. Porém, a natureza e a existência dos próprios pensamentos podem ser explicadas de uma forma que não introduz quaisquer propriedades não físicas especiais.⁷⁹ (CARRUTHERS, 2019, p. 121)

Segundo essa visão, não existiria uma lacuna ontológica entre propriedades e entidades físicas e não físicas, pois a realidade seria composta integralmente por propriedades e entidades físicas (isto é, um monismo ontológico). Haveria, na verdade, uma lacuna epistêmica, uma vez que os nossos conceitos sobre propriedades fenomenais seriam dualistas (isto é, um dualismo conceitual). De acordo com Carruthers (2019, pp. 123-125), os conceitos fenomenais estariam parcialmente isolados dos conceitos de terceira-pessoa (como os físicos, representacionais e funcionais). Poderíamos, portanto, nos referirmos aos nossos próprios estados conscientes, como a experiência do azul, sem empregarmos conceitos em terceira-pessoa (CARRUTHERS, 2020a). No entanto, mesmo que não tenham laços com conceitos de terceira-pessoa, os conceitos fenomenais ainda fariam referência a propriedades e entidades físicas. Em outras palavras, apesar de serem conceitualmente parcialmente isolados, os conceitos fenomenais permanecem se referindo a entidades físicas. Aqui especificamente, os conceitos fenomenais se referem a apenas conteúdos não conceituais globalmente compartilhados (CARRUTHERS, 2020a, pp. 8-9).⁸⁰

Por exemplo, quando penso no conteúdo fenomenal do azul da parede a minha frente, engajo-me em certos mecanismos cognitivos que geram o conceito fenomenal de “azul” (chamo de C_p). À primeira vista, no entanto, não há nada no meu uso de C_p em primeira-pessoa que se conecta diretamente com os processos neurofisiológicos que estão envolvidos na percepção consciente do azul. Ou seja, em razão dos conceitos fenomenais utilizados em primeira-pessoa serem parcialmente isolados, não haveria conexão direta entre o meu conceito

⁷⁹ “[...] not because our conscious states possess special properties of any sort (qualia), but rather from the distinctive first-person way in which it is possible for us to **think about** those states. It is because these thoughts lack conceptual connections with physical, functional, or even representational ones that we can conceive of the possibility of zombies; and it is also why we can see that even a complete physical/functional/intentional description of ourselves wouldn’t entail the truth of those thoughts. But the nature and existence of the thoughts themselves can be explained in a way that doesn’t introduce any special nonphysical properties, either”.

⁸⁰ “But our phenomenal concepts aren’t like that. We don’t use them with the intention of referring to whatever natural kind underlies our experiences [...]. On the contrary, we mean to refer just to qualities we are aware of in ourselves. (**These turn to be globally broadcast nonconceptual contents** [...].” (CARRUTHERS, 2019, pp. 147-148, ênfase minha)

fenomenal C_p e os conceitos em terceira-pessoa formulados nos termos neurofisiológicos envolvidos na percepção do azul.

Em suma, seria graças à essa ausência de conexões conceituais que surgiriam os equívocos dualistas atribuídos à consciência, como, por exemplo, a suposição de que a consciência é um fenômeno irreduzível, intrínseco, inefável e sobrenatural. Ademais, é também devido à falta dessas conexões que emergiriam os argumentos antifisicistas e a aparente incapacidade de descrevermos satisfatoriamente a consciência em termos puramente físicos, funcionais e/ou representacionais (CARRUTHERS, 2019; CARRUTHERS e VEILLET, 2007).

Claro, um dualista poderia admitir a existência de conceitos fenomenais, desde que eles fizessem referência a propriedades não físicas. No entanto, os fisicistas tipo-B asseguram que apesar de não possuírem conexões com os conceitos físicos os conceitos fenomenais se referem a *propriedades físicas* (uma vez que as propriedades fenomenais nada mais seriam do que propriedades físicas). De fato, por meio de experimentos com adultos e bebês de diferentes culturas ao redor do mundo, alguns psicólogos cognitivos sugerem que as nossas intuições dualistas derivam de *princípios conceituais dualistas inatos*. As nossas intuições dualistas sobre a consciência e outros estados mentais adviriam desses princípios (BERENT, 2020, 2021; BLOOM, 2004).⁸¹ Talvez esse seja um fator que corrobore empiricamente a hipótese da ECF. Os argumentos antifisicistas, neste sentido, estariam assentados em um conjunto de intuições dualistas inatas, embora todas as entidades mentais sejam puramente físicas. A origem e a função biológica dessas intuições, caso existam, ainda são incertas. Alguns autores sugerem que elas são subprodutos da nossa teoria da mente (BLOOM, 2004). Isso se contrastaria com a nossa “física intuitiva” (*intuitive physics*) que é inteiramente guiada por intuições fisicistas. Isso se tornaria claro em experimentos psicológicos sobre causalidade com bebês. Enquanto em experimentos envolvendo objetos os bebês esperam que um objeto *A* comece a se mover somente a partir do momento em que um objeto *B* o contacta, quando se trata de cenários

⁸¹ Isso não quer dizer que os pesquisadores aferiram a veracidade das hipóteses *metafísicas* do dualismo. O que os psicólogos intentaram argumentar é que possuímos princípios dualistas inatos que geram as intuições de cunho dualista. Por exemplo, em um experimento psicológico, pesquisadores reuniram dois grupos de crianças com o objetivo de descobrir se elas possuíam intuições dualistas em relação à transferência de mente de um corpo para outro. Um dos grupos participantes era composto de crianças ocidentais e o outro de crianças pertencentes a tribos indígenas. Interessante notar que a cultura dessas tribos indígenas considerava indesejável o emprego de estados mentais para explicar a causa do comportamento das pessoas. Em um dos cenários apresentados, os pesquisadores mostraram *implicitamente* às crianças uma criatura *A* (um pentágono) encarnando em uma criatura *B* (um triângulo) para atingir um objetivo (a obtenção de alimento). Foi observado que ambos os grupos de crianças tenderam à intuitivamente considerar que a mente de *A* saiu de seu corpo e, por sua vez, encarnou no corpo de *B* para obter o alimento (CHUDEK et al, 2018). Esse é um tipo de evidência que os proponentes da tese nativista dos princípios dualistas tipicamente levantam (BERENT, 2020, 2021; BLOOM, 2004; CARRUTHERS, 2020b).

envolvendo humanos esse tipo de expectativa por parte dos bebês deixa de existir (BERENT, 2020, pp. 63-64). No que diz respeito aos adultos, evidências de imageamento cerebral feitos por fMRI sugerem que as regiões cerebrais que são requisitadas para o raciocínio social, como a da teoria da mente, e para o raciocínio sobre causas físicas e mecânicas, como as relacionadas a objetos inanimados, são distintas. As áreas do córtex cingulado posterior e parietal medial, córtex pré-frontal medial, córtices temporal superior e parietal lateral inferior estão envolvidas no raciocínio social (JACK et al, 2013). Enquanto as áreas do córtex cingulado inferior que está conectado ao córtex pré-frontal lateral e parietal dorsal estão envolvidas no raciocínio sobre causas físicas e mecânicas (NAVARRO-CEBRIÁN e FISCHER, 2022). Foi observado que ao executar tarefas relacionadas ao raciocínio social, as atividades neuronais de suas regiões cerebrais correspondentes suprimiam a atividade neuronal de áreas que estão envolvidas com o raciocínio sobre causas físicas, e vice-versa (JACK et al, 2013). Isto não só sugere que as rotas neuronais envolvidas no raciocínio social e mecânico são diferentes, mas também que a ativação das áreas relacionadas ao primeiro raciocínio impede que as do segundo sejam ativadas, e vice-versa. Ora, isso forneceria uma evidência neuronal sobre o porquê a natureza da inferência que envolve a causalidade entre agentes é diferente da que envolve objetos inanimados, uma vez que áreas cerebrais distintas estão envolvidas em cada tipo de inferência e que a ativação de áreas envolvendo uma inferência suprime a atividade de áreas envolvendo a outra.

Quanto à natureza dos conceitos fenomenais, Carruthers (2019) identifica três propostas, são elas: os conceitos puramente de reconhecimento (CARRUTHERS, 2000), citacional (*quotational*) (PAPINEAU, 2002) e indicador (*indexical*) (CARRUTHERS, 2019, 2020a; PERRY, 2001; PRINZ, 2012).⁸² Apesar de inicialmente ter defendido a primeira concepção, atualmente Carruthers (2019, 2020a), seguindo e sofisticando a posição de Prinz (2012), sustenta a tese de que os conceitos fenomenais são indicadores baseados em relações de contato direto (*acquaintance-based indexical*) oriundos de conteúdos não conceituais globalmente compartilhados. Suponha, por exemplo, que neste momento estou percebendo, sob condições normais de luminosidade, a cor azul do copo situado a minha frente. Toda, ou boa parte, dessa informação chega em minha retina e é processada em meu sistema visual gerando o conteúdo perceptual “azul” (excluirei o conteúdo “copo” com o intuito de apenas simplificar o exemplo). Uma vez que este conteúdo não conceitual, o azul, adentra o espaço de trabalho

⁸² Devido ao escopo desta dissertação, me deterei somente à proposta dos indicadores baseados por contato. Sobre as outras concepções, cf. Carruthers, 2000; 2019, cap. 6; Papineau, 2002; 2020.

global e é amplamente compartilhado posso utilizá-lo de diversas formas e com diferentes objetivos. Ademais, quando este conteúdo perceptual adentra o meu espaço de trabalho global, torno-me apto para formar um novo conceito de modo especial, a saber, o conceito fenomenal “azul”. Segundo a concepção de Carruthers, o conceito fenomenal gerado pela minha experiência fenomenal do azul se dá sob a forma de um indicador (*indexical*) como, por exemplo, “*este-R*”. Ou seja, na medida em que o dado conteúdo não conceitual – a experiência do azul - adentra o espaço de trabalho global e se torna disponível aos sistemas executivos formuladores de juízos explícitos, ele pode ser conceitualizado sob a forma de um conceito fenomenal (CARRUTHERS, 2019, p. 123). Em resumo, o conceito fenomenal é gerado através do contato com a própria experiência correspondente.

Lembre-se de que para Carruthers, embora os conceitos fenomenais se refiram a conteúdos não conceituais globalmente compartilhados, eles são parcialmente isolados. Isto é, quando utilizados em primeira pessoa, os conceitos fenomenais não estão plenamente associados aos conceitos de terceira-pessoa. É justamente em virtude dessa falta de associação que surgiriam as nossas intuições dualistas acerca da consciência.

Considere os seguintes exemplos modificados de Carruthers (2019). Quando articulo pensamentos do tipo: “*este* é o tom de cor que gostaria para o meu carro”, ou “então *este-R* é afinal o que é como ver o azul”, os indicadores de cada um desses pensamentos introduzem conceitos fenomenais baseados no contato direto com a experiência consciente. No primeiro exemplo, se trataria de um conteúdo intencional não conceitual da cor que se tornou disponível aos sistemas executivos, resultando em um juízo de primeira ordem. O segundo exemplo, por sua vez, consistiria em um pensamento de ordem superior acerca do próprio conteúdo intencional não conceitual, a saber, o conteúdo perceptual gerado pela cor azul (CARRUTHERS, 2019, pp. 123-124). O que torna o indicador do segundo exemplo como sendo de ordem superior, segundo Carruthers (2019, p. 124), é a intenção ou disposição que o sujeito de experiências tem ao utilizar o conceito para se referir à experiência do azul e não a própria cor azul. Ora, ao utilizar o conceito “*este-R*” para se referir a experiência do azul, o sujeito está também implicitamente empregando o conceito “experiência”. É por essa razão que Carruthers considera que os conceitos fenomenais *são parcialmente* e não completamente isolados de outros conceitos.

No entanto, um dualista poderia contra-argumentar que as nossas intuições dualistas são de fato verdadeiras. As nossas intuições sobre eventos mentais e eventos físicos se diferem em razão das primeiras se referirem a propriedades sobrenaturais. Assim, as nossas intuições

dualistas emergem, pois de fato os conceitos fenomenais se referem a entidades não físicas. A falta de conexões conceituais só refletiria a relação existente entre os conceitos fenomenais e as suas referências (no caso, propriedades não físicas). Carruthers (2020a) considera essa possibilidade incabível. Como resposta, Carruthers fornece dois argumentos concisos que, segundo ele, também garantiriam a sua concepção de que os conceitos fenomenais se referem a propriedades físicas.

O primeiro argumento se apoia na chamada “simplicidade explanatória”. Segundo Carruthers (2020a, p. 9), a sua hipótese fisicista conseguiria explicar satisfatoriamente como surgem os argumentos antifisicistas. Bastaria, neste caso, que se apelasse à ausência de conexões entre os conceitos fenomenais e os conceitos físicos. Além disso, ela forneceria uma explicação mais simples e metafisicamente menos custosa do que a da hipótese dualista. Ora, na medida em que a primeira hipótese apela a somente propriedades físicas que, segundo Carruthers (2020a, p. 9), todos concordariam que existem, não haveria necessidade de postularmos propriedades especiais, como os *qualia*.⁸³ Ou seja, para que invocar entidades não físicas que são controversas se podemos manter em nossa hipótese propriedades físicas que são mais razoáveis e consensuais? Em resumo, “[...] a explicação mais simples é a de que a lacuna explanatória é uma *ilusão* produzida pela desconexão entre, de um lado, os nossos conceitos físicos em terceira-pessoa e os conceitos psicológicos, e, do outro, os nossos indicadores de ordem-superior em primeira-pessoa.”⁸⁴ (CARRUTHERS, 2020a, p. 9, ênfase minha). Ao fim e ao cabo, Carruthers está utilizando nada mais que a navalha de Ockham, ou seja, se for possível explicar com sucesso um determinado evento ou fenômeno sem que seja necessário a postulação de propriedades adicionais, faça-o.⁸⁵

⁸³ Dentre as propriedades consensuais Carruthers (2020a, p. 9) cita: o compartilhamento global, o conteúdo representacional, os conceitos do tipo indicadores e os de ordem superior. Todas são debatíveis em maior ou menor grau. No entanto, o conteúdo representacional sem dúvida é a mais polêmica dentre elas, figurando no centro de um fervoroso debate na filosofia da mente e nas ciências cognitivas, a chamada “Guerras de Representação”. Embora as noções de “representação” e “conteúdo representacional” ainda façam parte da ortodoxia das ciências da mente, desde o final da década de 90 elas vêm sendo atacadas pelos proponentes das teses corporificadas. Com base nos desenvolvimentos da psicologia e das neurociências, alguns autores defendem que tais propriedades sequer existem (RAMSEY, 2007). Assim, ao contrário do que Carruthers afirma, “conteúdo representacional” e “representação” são entidades polêmicas, embora façam parte do arcabouço ortodoxo das ciências da mente. Cf. Pereira et al, no prelo; Shapiro e Spaulding, 2021.

⁸⁴ “[...] *the simplest explanation is that the explanatory gap is an illusion produced by a disconnect between our third-person physical and psychological, on the one hand, and our first-personal higher-order indexical ones, on the other.*”

⁸⁵ Alguns autores argumentam que nem sempre a teoria mais simples é de fato a verdadeira ou a melhor dentre as suas concorrentes. Eles argumentam que a própria “simplicidade” poderia ser substituída por outros critérios, caso ela se demonstre a menos adequada para tratar sobre o assunto em questão. Chalmers (2022, pp. 76-79), por exemplo, considera que o apelo à simplicidade não prova a tese de que o mundo externo existe, o que torna plausível a hipótese de a realidade ser uma simulação virtual. Sobre a hipótese de que podemos estar vivendo em uma realidade simulada, cf. Bostrom, 2003; Chalmers, 2022.

O segundo argumento diz respeito à implementação da ECF contra os argumentos antifísicistas. Embora Carruthers utilize a ECF para rebater vários desses argumentos, aqui me restringirei a apenas dois deles. Serão respectivamente o argumento da concebibilidade dos zumbis e o argumento do quarto da cientista Mary (CHALMERS, 1996; JACKSON, 1986). Segundo Carruthers (2019, p. 126), uma vez que os conceitos fenomenais são essencialmente de primeira-pessoa, os seus usos não implicariam em conceitos físicos, pois os primeiros são parcialmente isolados dos segundos. Por essa razão, segundo Carruthers (2019, p. 126), é plenamente coerente a formulação de juízos que concebem os zumbis, como: “Poderia ser o caso de haver um ser que é fisicamente, funcionalmente e representacionalmente idêntico a mim, mas que não possuísse *este-E*”⁸⁶ (CARRUTHERS, 2019, p. 126) (onde “*este-E*” é um indicador que se refere à minha experiência consciente). Ou seja, ao formular o juízo acima, não haveria nada em meu uso de “*este-E*” que se conectaria com conceitos físicos, tais como “neurônio” e “cérebro”. Com efeito, apesar dos zumbis serem concebíveis, nenhuma propriedade especial e não física seria derivada desse juízo. O indicador continuaria se referindo a propriedades físicas, no caso os conteúdos não conceituais globalmente compartilhados. Em suma, para Carruthers (2019, p. 126), a concebibilidade dos zumbis é explicada devida a falta de conexões entre os conceitos fenomenais e os conceitos em terceira-pessoa, cuja aplicação dos primeiros conceitos não implica na dos segundos.

Até mesmo os zumbis, segundo Carruthers (2020a, p. 9), seriam capazes de conceberem outros zumbis, dada a falta das conexões conceituais supracitadas. Os zumbis conseguiriam articular pensamentos como: “[p]oderia haver uma criatura exatamente como eu em todos os aspectos físicos, funcionais e representacionais, mas que não possuíssem nada como *isto*”⁸⁷ e “[n]ão importa o quanto eu conheça sobre os fatos físicos, funcionais e representacionais envolvidos na percepção, isso ainda não explica o porquê desses fatos darem surgimento a estados como *este*”⁸⁸ (onde os indicadores se referem a compartilhamento global de conteúdos não conceituais que os zumbis não possuiriam) (CARRUTHERS, 2020a, p. 9). Novamente, para Carruthers tudo isso poderia ser explicado através da falta de conexões entre conceitos fenomenais e físicos, e não devido à existência de propriedades não físicas. O problema emergiria do esquema conceitual, e não devido a propriedades sobrenaturais.

⁸⁶ “*There could be a being who is physical, functional, and representational duplicate of myself, who nevertheless lacks **this-E**.*”

⁸⁷ “[t]here could be a creature exactly like myself in all physical, functional, and representational respects who nevertheless lacks anything like **this**.”

⁸⁸ “[n]o matter how much I know about the physical, functional, and representational facts involved in perception, that still won’t explain why those facts give rise to states like **this**.”

No que diz respeito ao quarto de Mary, Carruthers (2019, p. 127) afirma que ao sair de seu aposento inteiramente preto e branco e experienciar cores pela primeira vez, a cientista de fato adquiriria um novo conhecimento. Ao experienciar, por exemplo, o céu azul ela aprenderia proposições do tipo: “então *este-R* é o que as pessoas experienciam quando veem o céu azul” (onde “*este-R*” se refere à sua experiência de perceber o céu azul). Segundo Carruthers, esta seria uma proposição que Mary só poderia aprender por meio dessa experiência, uma vez que a proposição envolve um conceito fenomenal que é “[...] fundamentada no contato com representações não conceituais [...]”⁸⁹ (CARRUTHERS, 2019, p. 127). Deste modo, ao ter pela primeira vez esses conteúdos não conceituais sendo globalmente compartilhados em sua mente, Mary acessaria um novo conhecimento proposicional e um novo pensamento verdadeiro, que em seu quarto não poderia conhecer. Entretanto, para Carruthers (2019, p. 127), Mary permaneceria tendo conhecimento sobre todos os *fatos* relativos ao pensamento e à proposição supracitada. Isto é, ela conheceria todos os fatos que envolvem o compartilhamento dessas informações em sua mente quando se percebe a cor azul. A novidade para Mary estaria, portanto, “[...] *no nível do pensamento*, resultante das novas representações não conceituais que foram pela primeira vez transmitidas em sua mente (conteúdo não conceitual de cor).”⁹⁰ (CARRUTHERS, 2019, p. 127, ênfase minha). Segundo Carruthers (2019, p. 127), novamente não foi necessário postularmos propriedades não físicas para a explicação deste experimento de pensamento. Tudo seria explicado através de termos físicos, como o do compartilhamento global de conteúdo não conceitual na mente.

O que afinal se extrai desses argumentos, segundo Carruthers (2019, 2020a), é o seguinte: o emprego dos conceitos fenomenais do tipo indicador baseado em um contato direto com a experiência responderia satisfatoriamente as objeções antifisicistas, pois é devido à falta de conexões entre os conceitos fenomenais e os de terceira-pessoa que emergem as intuições dualistas dessas objeções, e não em razão da existência de propriedades não físicas. Ou, em suas próprias palavras: “[...] nós não precisaríamos apelar a nada além de conceitos fenomenais e conteúdos compartilhados como forma de explicar como esses experimentos de pensamento emergem”⁹¹ (CARRUTHERS, 2020a, p. 9). Seria deste modo que a ECF forneceria o ingrediente faltante para que a GWT enfim se tornasse a teoria completamente reducionista da consciência.

⁸⁹ “[...] *grounded in acquaintance with nonconceptual representations [...]*”

⁹⁰ “[...] *at the level of thought, resulting from new nonconceptual representations being broadcast in her mind for the first time (nonconceptual color content).*”

⁹¹ “[...] *we don't need to appeal to anything other than phenomenal concepts and broadcast contents in order to explain how those thought experiments arise.*”

Na próxima seção apresentarei aquilo o que motiva a conclusão cética de Carruthers segundo a qual não há questão de fato sobre se os animais possuem consciência, também nomeado por Birch (2020b) de “o condicional surpreendente de Carruthers”. Em seguida, exporei os dois argumentos centrais que, segundo Carruthers (2019, pp. 153-161), justificariam esse condicional.

2.5 A Motivação por trás da Postura Cética de Carruthers

O condicional surpreendente de Carruthers, que resulta em seu ceticismo, é expresso da seguinte forma por Birch:

*Se a GWT é a teoria correta e totalmente reducionista da consciência fenomenal em humanos, então, para todos os animais não humanos, não há questão de fato sobre se eles possuem ou não estados fenomenalmente conscientes.*⁹² (BIRCH, 2020b, p. 7)

Ou nas palavras do próprio Carruthers:

[S]e a teoria do espaço de trabalho global da consciência fenomenal está correta, e é completamente reducionista por natureza, então devemos parar de fazer perguntas sobre a consciência em animais não humanos – não porque essas perguntas são muito difíceis para responder, mas porque não existem fatos substantivos a serem descobertos.⁹³ (CARRUTHERS, 2018b, p. 47)

Como Birch (2020b) bem salienta, o condicional que Carruthers fornece consiste no seguinte *modus ponens*:

- A) Se a GWT é a teoria correta e completamente reducionista da consciência fenomenal, então não há questão de fato acerca da atribuição de consciência em animais não humanos.
- B) A GWT é a teoria correta e completamente reducionista da consciência.

⁹² *If GWT is a correct and fully reductive theory of phenomenal consciousness in humans, then, for all non-humans, there is no fact of the matter as to whether or not they possess phenomenally conscious states.*

⁹³ “[I]f a global workspace theory of phenomenal consciousness is correct, and is fully reductive in nature, then we should stop asking questions about consciousness in nonhuman animals -not because those questions are too hard to answer, but because there are no substantive facts to discover.”

C) Portanto, não há questão de fato acerca da atribuição de consciência em animais não humanos.

O que motivaria o argumento radical apresentado logo acima? Segundo Carruthers (2019, p. 143), ele é motivado pelo descompasso (*mismatch*) existente entre de um lado a natureza tudo-ou-nada da consciência fenomenal e o espaço de trabalho global dos humanos e de outro o espaço de trabalho global dos animais. De acordo com o que se viu, a GWT supostamente seria a teoria correta e completamente reducionista da consciência em humanos. Segundo a interpretação que Carruthers faz da GWT, a consciência nada mais seria do que conteúdos não conceituais globalmente compartilhados. Ademais, a consciência consistiria em um fenômeno tudo-ou-nada, isto é, ou o sujeito está consciente ou não está. A consciência, portanto, não admitiria gradações.

Por outro lado, quando se investiga os outros animais, observa-se diferentes formas de espaço de trabalho global em arquiteturas cognitivas distintas (ainda que em certa medida comparáveis às humanas). Nenhum outro animal conhecido possui, por exemplo, o sistema consumidor tal como o do relato verbal e poucos possuem o da metacognição, do planejamento, do raciocínio e o da teoria da mente. Mesmo os sistemas compartilhados entre humanos e animais admitem variação de complexidade. Por exemplo, a nossa capacidade de refletir acerca dos nossos próprios pensamentos é, provavelmente, mais complexa do que a de nossos parentes mais próximos, como os chimpanzés, e a dos nossos parentes distantes, como os chocos. Em suma, o que se observa nas diferentes linhagens são arranjos variados de espaços de trabalho global, sendo uns mais parecidos com o dos humanos (como o dos grandes primatas) e outros completamente distintos (como o dos cefalópodes).

Como se pode notar, existe um descompasso entre a consciência, que é um fenômeno tudo-ou-nada e completamente explicada pela GWT em humanos, e o espaço de trabalho global dos outros animais que admite gradações. Ora, uma vez que se rejeite a tese gradualista da consciência, não se pode afirmar que os animais, a depender da sua arquitetura cognitiva, possuem graus diferentes de consciência, sendo uns mais complexos que os outros. Portanto, ou se expressa que os estados *E* apresentados pelos animais são conscientes ou se expressa que *E* não consiste em um estado consciente, pois caso o contrário existiria uma contradição (CARRUTHERS, 2019, p. 143). O descompasso possui a seguinte estrutura:

A) A consciência é um fenômeno tudo-ou-nada ou gradualista.

B) Se consciência é um fenômeno tudo-ou-nada, então ela não é um fenômeno gradualista.

C) A consciência é um fenômeno tudo-ou-nada.

D) Logo, a consciência não é um fenômeno gradualista.

Assumindo que a consciência seja apenas conteúdo não conceitual globalmente compartilhado:

A) A consciência é um fenômeno tudo-ou-nada.

B) Em humanos o espaço de trabalho global é um fenômeno tudo-ou-nada.

C) Em animais não humanos o espaço de trabalho global admite gradações.

D) Logo, a consciência é um fenômeno tudo-ou-nada em humanos e gradualista em animais não humanos.

Carruthers (2019, pp. 143-144) admite duas respostas possíveis ao descompasso. A primeira delas é a de que devemos rejeitar a GWT. A existência do descompasso seria um indicativo de que a GWT não consiste na teoria completamente reducionista da consciência e, por consequência, devemos abandoná-la. Carruthers, contudo, insiste que essa não é a melhor opção, pois “[...] não existiria um descompasso entre a nossa concepção de consciência fenomenal e o compartilhamento global no caso dos humanos”⁹⁴ (CARRUTHERS, 2019, p. 143). Segundo o autor, a GWT explicaria todos os fatos sobre a consciência em humanos, sem conflitar com o nosso uso em primeira pessoa dos conceitos fenomenais e da “consciência” (CARRUTHERS, 2019, pp. 143-144). Ademais, a dissolução das objeções antifisicistas através da combinação entre a GWT e a ECF consistiria em uma forte evidência a favor da GWT. Portanto, como nenhum fato sobre a consciência humana fugiria do escopo da GWT e a ECF fornece a justificativa para que ela seja de fato uma teoria completamente reducionista, então não haveria razões para rejeitá-la.

No entanto, poder-se-ia argumentar que, segundo as nossas intuições, ao menos os animais filogeneticamente próximos de nós seriam conscientes. Ou seja, como geralmente

⁹⁴ “[...] *there isn't any mismatch between our conception of phenomenal consciousness and global broadcasting in human case.*”

consideramos intuitivo que animais como os chimpanzés, canídeos e outros mamíferos sejam conscientes, provavelmente eles de fato o são. Ora, não seria evidente que chimpanzés experienciam estados conscientes como os de dor e prazer? Carruthers (2019, p. 144), porém, considera um equívoco o emprego de intuições para solucionar o descompasso. Os motivos seriam os seguintes: a) esse tipo de intuição não faz parte das condições de posse de um conceito fenomenal (o que não é igual a afirmar que a posse de conceitos fenomenais é uma condição necessária ou suficiente para o organismo ser consciente) e b) alguém poderia possuir um conceito fenomenal e atribuir consciência a todos ou mesmo a nenhum dos animais. Pense, por exemplo, em um pampsiquista e um ilusionista radical respectivamente. No caso do pampsiquista, ele “[...] aparenta não conter restrições conceituais relativas à extensão do conceito [“consciência”]”⁹⁵ (CARRUTHERS, 2019, p. 144, acréscimo meu). Em outras palavras, a extensão dos conceitos fenomenais de um pampsiquista seria ubíqua. Enquanto para um ilusionista radical nenhum animal seria consciente, incluindo os humanos, uma vez que propriedades fenomenais não existem. Assim, a extensão dos conceitos fenomenais de um ilusionista radical seria vazia. Portanto, o uso das intuições não parece, segundo Carruthers, consistir no melhor guia para a de atribuição de consciência a outros animais, pois elas variam de acordo com a posição filosófica adotada.

É curioso notar que embora Carruthers rejeite o uso de intuições para solucionar o descompasso, ele as utiliza largamente na sustentação da tese tudo-ou-nada da consciência. Isso fica claro na medida em que ele utiliza a intuição de que é difícil conceber casos limítrofes de consciência para rejeitar a posição gradualista (CARRUTHERS 2019, pp. 21-22). No entanto, cada vez mais autores desafiam essas intuições através de argumentos filosóficos robustos e evidências empíricas (BIRCH et al, 2020; COHEN et al, 2022; GODFREY-SMITH, 2020a; HALL, 2022; TYE, 2021; VEIT, 2022b). Mais do que isso, alguns autores argumentam que a existência ou concebibilidade de casos limítrofes não é uma condição necessária para que se defenda o gradualismo, o que torna ainda mais frágil a objeção de Carruthers (BIRCH, 2021). Tendo isso em vista, me parece ser irrazoável o apelo a intuições como forma de determinar a natureza da consciência.

A segunda resposta, a qual Carruthers defende, é a de que não há questão de fato acerca da consciência em animais (CARRUTHERS, 2018a, 2018b, 2019, 2020a). Para além dos humanos, segundo o autor, existiriam apenas casos de “*indeterminação semântica*”. Mas o que afinal é indeterminação semântica? Carruthers não a define explicitamente e tampouco a

⁹⁵ “[...] would appear to be no conceptual constraints on the extension of the concept.”

distingue de outros tipos de indeterminação. Como forma de tornar mais claro, vejamos brevemente três possíveis tipos de indeterminação: a epistêmica, a metafísica e a semântica.⁹⁶

Por indeterminação epistêmica entendo aquela onde um agente não privilegiado⁹⁷ R é incapaz de acessar o conhecimento sobre p devido a alguma razão ou conjunto de razões Q . Assim, R manteria uma relação de indeterminação epistêmica com p quando não se pode afirmar que: ou R conhece p devido a Q ou R não conhece p devido a Q . Note que R não se encontra meramente em um estado de ignorância, mas em um estado cujo conhecimento ou não de p é indeterminado. A indeterminação epistêmica pode assumir ao menos duas versões: uma moderada e outra robusta. Segundo a moderada, o conhecimento sobre p está epistemicamente inacessível para R por insuficiência de informações relevantes e/ou pelo estágio tecnológico precário sobre p . Por exemplo, devido a questões tecnológicas não podemos atualmente obter conhecimento direto sobre como se deu o Big Bang. No entanto, isso não quer dizer que somos totalmente ignorantes sobre o assunto, tendo em vista que dispomos de modelos teóricos e evidências indiretas, mesmo que rudimentares, sobre o início do cosmos. A versão robusta, por sua vez, considera que o conhecimento sobre p é *por princípio* inacessível a R . Ou seja, R é incapaz de conhecer p independentemente de quaisquer motivos, sejam eles de natureza tecnológica, teórica ou pela indisponibilidade de informações e evidências relevantes sobre p . Por exemplo, certas posições kantianas consideram ser intrinsecamente impossível para o ser humano a obtenção do conhecimento sobre a coisa em si (*noúmeno*). Os *misterianistas*, de forma semelhante, argumentam que a consciência é um mistério que nunca será resolvido, pois a nossa inteligência não foi feita para conhecê-la (MCGINN, 1999).

A indeterminação metafísica, por sua vez, diz respeito à própria constituição da realidade. Ela independe da forma como teorizamos, das evidências relevantes disponíveis ou das tecnologias que dispomos para investigar o mundo. Grosso modo, essa concepção afirma que existiriam entidades e propriedades da nossa realidade que são indeterminadas, ou vagas, devido aos próprios aspectos que as constituem. Por exemplo, embora existam partes que definitivamente pertencem a montanha Kilimanjaro e outras que não a pertencem, alguns autores argumentam que é indefinido quais são as partes exatas que a constituem. Até mesmo o espaço exato em que a montanha se localiza, isto é, os seus limites, seriam vagos por natureza. De modo semelhante, alguns autores argumentam que a identidade pessoal de um indivíduo

⁹⁶ O intuito aqui é fornecer apenas uma operacionalização dos diferentes tipos de indeterminação, algo que Carruthers não o faz. Como o assunto é complexo e controverso, sugiro a seguinte bibliografia para obter informações refinadas: Taylor e Burgess, 2015; Williams, 2008.

⁹⁷ Excluo, portanto, as entidades oniscientes.

também seria metafisicamente indeterminada, uma vez que é vago o momento exato em que ela surge ou desaparece. Por exemplo, seria constitutivamente indefinido o momento exato em que um indivíduo *P* falece por um motivo *Q* (e. g. acidente fatal e doenças) (WILLIAMS, 2008). Também seria indeterminado o momento exato em que uma população de indivíduos da espécie *P* torna-se uma espécie *Q*. Todos os exemplos supracitados supostamente possuiriam uma indeterminação nos aspectos que metafisicamente os constituem. Existe uma grande controvérsia se de fato existem entidades vagas. Evans (1978), por exemplo, argumenta que devido a inconsistências lógicas envolvendo a Lei de Leibniz⁹⁸ não poderiam existir entidades vagas. Por fugir do escopo aqui delimitado não me debruçarei sobre esse debate.

Por fim, a indeterminação semântica, em linhas gerais, envolveria as noções típicas da linguagem, como: significado, referência, conceitos, conteúdo, representação dentre outras. Como tudo o que envolve o terreno da linguagem, a noção de “indeterminação semântica” também é movediça e de difícil caracterização (TAYLOR e BURGESS, 2015). Assim, me deterei apenas aos usos, mesmo que idiossincráticos, adotados por Carruthers. Ao que tudo indica, Carruthers (2019) entende indeterminação semântica de duas maneiras distintas. A primeira envolve a referência de um conceito e a segunda o valor de verdade de um certo juízo. Segundo a primeira, teríamos indeterminação quando a referência de um dado conceito é vaga. Isto é, quando o uso de um conceito *C* não consegue rastrear com sucesso a sua referência. O segundo tipo de indeterminação semântica diz respeito aos juízos que emitimos sobre *p* que, por algum motivo, são destituídos de valor de verdade, ou seja, não são verdadeiros e nem falsos.⁹⁹ Por exemplo, certas teses não cognitivistas em metaética sustentam que por serem emoções, os juízos morais que emitimos são destituídos de valor de verdade, não sendo verdadeiros e nem falsos (MILLER, 2013). Realizado as distinções, vejamos a seguir os dois argumentos que embasam a posição radical de Carruthers.

⁹⁸ Geralmente formalizada como $\forall F (Fx \leftrightarrow Fy) \rightarrow x = y$. Lê-se como se para toda propriedade *F*, um certo objeto *x* possui *F* se e somente se um objeto *y* também possui *F*, logo os objetos *x* e *y* são idênticos.

⁹⁹ Birch (2022a) considera que a posição fisicista tipo-B de Papineau (2002) enfrenta uma indeterminação ainda mais severa do que a de Carruthers (2019). Para Papineau, um conceito fenomenal *C_n* rastreia com sucesso dois níveis de organização diferentes, a saber, uma propriedade funcional de nível superior *F* e o seu realizador neuronal *G*. Por exemplo, o processamento de informação visual (*F*) e a atividade no córtex V1 (*G*). Deste modo, não haveria como saber se *C_n* se refere a *F* ou *G*, não em razão de uma limitação epistêmica e/ou tecnológica, mas por motivações semânticas. Em outras palavras, haveria uma indeterminação intrinsecamente semântica que nos impossibilitaria de saber se *C_n* se refere a *F* ou *G*.

2.6 O Argumento Semântico Negativo

De forma geral, o argumento semântico negativo expressa o seguinte. Não existiria nada em nossas *intenções* quando utilizamos os conceitos fenomenais que indique a semelhança ou dissemelhança entre a arquitetura cognitiva dos humanos e a dos animais que nos possibilite nos pronunciar sobre a consciência animal. Em outras palavras, não há nada em nosso *uso em primeira-pessoa* dos conceitos fenomenais que fixe a sua extensão para além dos organismos cuja arquitetura é relevantemente similar a nossa (CARRUTHERS, 2019, p. 153). Antes de prosseguir, é importante salientar que Carruthers (2019, p. 125) adota uma semântica que chamarei de “*intencional*”, onde o que fixa a extensão de um conceito fenomenal C_n é a *intenção* do sujeito possuidor de C_n . Por “*intencional*” entendo a articulação deliberadamente explícita do sujeito de experiências para se referir a x ao utilizar C_n .¹⁰⁰ No caso de Carruthers, a variável *sempre* será preenchida por algum conteúdo não conceitual globalmente compartilhado. Os motivos pelos quais Carruthers (2019, pp. 125-126) endossa tal semântica são os seguintes. Primeiro, como um C_n é fundamentalmente caracterizado como sendo de primeira-pessoa, para a fixação de sua extensão *não seria possível apelar a deferências de outros indivíduos*, como um especialista, tampouco recorrer a um livro-texto sobre o assunto (BURGE, 1979). Além disso, os C_n não consistiriam em conceitos de tipo natural, uma vez que a sua referência não é determinada por nenhuma suposta essência subjacente (como ocorreria, por exemplo, com o termo “*água*” cuja extensão é determinada pelo composto químico H_2O).¹⁰¹

Empregaríamos C_n com apenas a intenção de nos referirmos ao conteúdo de nossas próprias experiências conscientes. Assumindo o irrealismo de *qualia*, poder-se-ia dizer que “[...] as extensões dos conceitos fenomenais são fixadas por suas disposições classificatórias ao empregá-los.”¹⁰² (CARRUTHERS, p. 125). Ou seja, quando articulo um pensamento com o conteúdo “*esta é uma experiência prazerosa*”, a extensão do indicador “*esta*” será “[...] o conjunto de conteúdos não conceituais que evocaria as mesmas disposições-para-julgar como as que subjazem o pensamento em questão (ou algo do tipo).”¹⁰³ (CARRUTHERS, 2019, pp. 125-126, ênfase minha). Note que, para Carruthers, os conceitos fenomenais se referem

¹⁰⁰ Lembre-se de que para Carruthers (2019, p. 123) os conceitos fenomenais só podem ser utilizados *explicitamente*, isto é, após o conteúdo não conceitual estar disponível aos sistemas executivos.

¹⁰¹ Bayne e Shea (2020) discordam deste ponto. Os dois filósofos defendem que a consciência pode ser um tipo natural, porém não no sentido de ser uma propriedade única, eterna e imutável, mas um conjunto de propriedades homeostáticas. Ademais, os autores consideram que o termo “consciência” pode ser ao mesmo tempo um conceito fenomenal e natural dito “*não canônico*”.

¹⁰² “[...] *the extensions of one’s phenomenal concepts get fixed by one’s classificatory dispositions when employing them.*”

¹⁰³ “[...] *as the set of non-conceptual contents that would evoke the same dispositions-to-judge as those that underlie the thought in question (or something of the sort).*”

diretamente à própria experiência, que nada mais é que conteúdo não conceitual globalmente compartilhado.

Retomando o argumento semântico negativo, nada em nosso uso dos conceitos fenomenais fixaria a extensão para os estados mentais de outros animais. Por exemplo, quando percebo conscientemente uma pintura azul diante do meu campo de visão sob condições normais de luminosidade, posso pensar algo como: “*esta* é uma bela experiência”. Neste caso, não existiria nada embutido no indicador baseado por contato “*esta*” que fixaria a sua extensão para arquiteturas cognitivas para além da minha própria. Quando faço uso em primeira-pessoa de “*este-R*” ou qualquer outro conceito fenomenal desse tipo, a minha intenção só fixaria a extensão para o contexto em primeira-pessoa em que ele está inserido, a saber, o determinado conteúdo não conceitual globalmente compartilhados aos sistemas consumidores relacionados ao meu estado perceptual consciente em questão.

Nesse sentido, existiria apenas uma maior ou menor similaridade entre o espaço multidimensional dos animais e o compartilhamento global de informações no espaço de trabalho dos humanos. Como a consciência não é um fenômeno que admite graus, deveria haver uma linha demarcatória que separasse os animais conscientes dos inconscientes à medida em que as mentes deles se tornassem menos semelhantes à nossa. Ora, como não há nada em nossos usos dos conceitos fenomenais que fixam esses limites, por conseguinte, não existiria nada para ser descoberto no que tange à consciência dos animais. Logo, não existiria questão de fato envolvendo a consciência em animais.¹⁰⁴

Como foi visto em parágrafos anteriores, Carruthers não poderia apelar a propriedades como *qualia* ou pintura mental como sendo os referentes dos nossos conceitos fenomenais. A admissão de tais propriedades acarretaria o compromisso com o realismo de *qualia*, posição que ele rejeita. Deste modo, conclui-se que não existiria questão de fato sobre se os outros animais são conscientes. Note que o argumento semântico negativo *não* afirma que os animais *não são* conscientes, mas antes que não há fato em razão de uma indeterminação semântica. Isto é, não há nada em nosso uso dos conceitos fenomenais que fixa “[...] o quão fisicamente ou funcionalmente similar um animal deve ser ao caso humano para que seja qualificado como consciente”¹⁰⁵ (CARRUTHERS, 2019, p. 154).

¹⁰⁴ “*In fact, there is nothing about our first-person concept of phenomenal consciousness that fixes its extension across the mental states of other species. Hence there is no fact of the matter whether members of any those species have phenomenally conscious states or not.*” (CARRUTHERS, 2019, p. 153)

¹⁰⁵ “[...] *how physically or functionally similar an animal must be to the human case to qualify as conscious.*”

Carruthers também faz questão de sublinhar que este argumento não se assenta na noção de “vagueza”. Ou seja, a conclusão de que não existe questão de fato se animais possuem ou não consciência não é devido a “[...] seus estados caírem em algum lugar no limite vago de nosso conceito de consciência”¹⁰⁶ (CARRUTHERS, 2019, p. 153), uma vez que “consciência” não admite casos limítrofes. Ora, a admissão de tais casos implicaria no compromisso com alguma forma de gradualismo, tese que Carruthers rejeita. Nesse sentido, não haveria indeterminação metafísica no que concerne à consciência, pois o conceito “consciência” e a própria consciência seriam, segundo Carruthers, fenômenos tudo-ou-nada.

Ademais, o argumento não se assenta em uma indeterminação epistêmica, pois não é graças à nossa limitação cognitiva, tecnológica ou das evidências disponíveis que não podemos saber se os animais possuem ou não consciência. Mesmo se conhecêssemos tudo sobre os aspectos de terceira-pessoa de uma abelha, por exemplo, nada de novo seria acrescentado sobre se ela possui ou não consciência. O motivo para isso é que o nosso conceito fenomenal só fixa o seu referente para organismos com arquiteturas cognitivas relevantemente similares a nossa. Ora, como não há nada em nossas intenções quando utilizamos um conceito fenomenal que fixa o quão similar conosco os estados mentais da abelha devem ser para que sejam conscientes, então não existe questão de fato sobre se elas possuem ou não consciência (CARRUTHERS, 2019, p. 155).

2.7 O Argumento Semântico Positivo

Diferentemente do argumento anterior, o argumento semântico positivo independe da premissa de que a consciência é um fenômeno tudo-ou-nada. Embora também se assente na premissa de que a consciência é apenas conteúdo não conceitual globalmente compartilhado, o argumento diz respeito à semântica dos juízos que atribuem consciência aos animais. Particularmente abordando as condições de verdade que envolvem tais juízos. Como Carruthers sublinha, deve-se lembrar que “consciência” é um conceito fundamentalmente de primeira-pessoa, “[...] fundado nas próprias capacidades de pensamento acerca de seus estados perceptuais compartilhados globalmente”¹⁰⁷ (CARRUTHERS, 2019, p. 155). Por exemplo, quando contemplo a pintura *No. 61 (Rust and Blue)* do pintor norte-americano Mark Rothko, posso pensar algo como: “então *isto* é como ver um azul vívido”, onde o indicador se refere tão somente a minha experiência do azul vívido tal como o apresentado pela pintura. Como vimos,

¹⁰⁶ “[...] *their states fall somewhere within the vague boundary of our concept of consciousness.*”

¹⁰⁷ “[...] *grounded in one’s capacities for thought about one’s own globally broadcast perceptual states.*”

para Carruthers essa experiência consiste apenas em conteúdo não conceitual globalmente compartilhado que se tornou disponível aos diversos sistemas consumidores, como os de pensamento de ordem superior, da percepção e da linguagem que estão envolvidos no meu pensamento acerca da pintura de Rothko.

Poderíamos então nos perguntar quais seriam as condições de verdade de juízos como o do tipo: “o animal *S* possui estados perceptuais como este-*R*” (CARRUTHERS, 2019, p. 156). Algumas opções são sumariamente rejeitadas por Carruthers (2019, 2020a). A primeira delas consiste na hipótese de que os nossos conceitos fenomenais de alguma maneira se referem a tipos naturais. A razão pelo qual se deve rejeitá-la é, segundo Carruthers (2019), a de que quando utilizamos um conceito fenomenal não o utilizamos com a *intenção* de que ele seja um conceito de tipo natural ou de que ele rastreie uma essência que supostamente subjaz a nossa experiência.¹⁰⁸ Novamente observa-se o apelo de Carruthers à semântica intencional como o fixador da referência de um conceito fenomenal. A segunda opção é a de que o conceito fenomenal se refere a propriedades fenomenais, como os *qualia*, que estariam presentes nos estados mentais dos animais, incluindo o dos humanos. Como já foi visto, por Carruthers negar que existam tais propriedades, elimina-se essa opção. A terceira opção, por sua vez, é a de que os conceitos fenomenais se refeririam a apenas conteúdos não conceituais. Contudo, mesmo estados perceptuais inconscientes possuem conteúdo não conceitual (pense, por exemplo, nos casos de visão cega). Ora, para que um conteúdo não conceitual se torne consciente, segundo Carruthers, ele deve adentrar o espaço de trabalho global e ser amplamente compartilhado. Portanto, o mero fato de ser um conteúdo não conceitual não garante o valor de verdade dos juízos, dado que ele *não* satisfaz o requisito de *ser globalmente compartilhado*. Ao fim e ao cabo, Carruthers considera que o que fixaria o valor de verdade de juízos que atribuem consciência é o seguinte contrafactual:

A criatura *C* possui estados perceptuais que são este-*E*. [...] Se as disposições para julgar que subjazem o meu uso do conceito ESTE-*E* fossem instanciadas na criatura *C*, então elas emitiriam um julgamento de que alguns dos estados perceptuais da criatura são este-*E*.¹⁰⁹ (CARRUTHERS, 2019, p. 157)

¹⁰⁸ Carruthers se inspira nos argumentos de Kripke (2018) contra a posição de que a consciência é um tipo natural. Para maiores detalhes cf. Carruthers, 2019, cap. 7, seção, 7.2.

¹⁰⁹ “*Creature C has perceptual states that are this-E. [...] If the dispositions-to-judge that underlie my use of the concept THIS-E were to be instantiated in creature C, then they would issue in a judgment that some of the creature’s perceptual states are this-E.*”

Como forma de tornar claro o que Carruthers tem em mente, consideremos a seguinte analogia (CARRUTHERS, 2019, p. 157). Considere que João, um ser humano neurotípico, está prestes a se mudar. Suponha que João esteja andando em um bairro *B* localizado em uma cidade *C*, que nunca tivera visitado. Em uma de suas andanças por *B*, João faz a seguinte afirmação a Gabriela, sua esposa: “*Este* é o tipo de local no qual devemos viver!” (CARRUTHERS, 2019, p. 157). Como Carruthers aponta, o referente de “*este* tipo de local” não são objetos singulares como as casas ou os objetos que decoram o bairro. O demonstrativo em questão tampouco se refere a uma essência que subjaz *B*, onde *B* seria um tipo natural. Suponha agora o seguinte cenário: em uma andança sob as mesmas condições externas (e. g., clima) e internas (e. g., temperamento) por *C*, mas no bairro *D*, Gabriela indaga a João: “Este bairro é *daquele* tipo?” (CARRUTHERS, 2019, p. 157). O problema é: caso João respondesse “sim” ou “não”, o que fixaria o valor de verdade do juízo enunciado? Carruthers responde o seguinte:

[S]e alguém visitasse o outro bairro sob as mesmas condições com os mesmos precursores (criando então o mesmo conjunto de disposições para julgar) então ele contaria (ou não) aquele outro bairro como o mesmo no aspecto relevante. Ou, de forma um pouco mais abstrata: se alguém com exatamente as mesmas disposições para julgar que estavam envolvidas no juízo sobre o primeiro bairro fosse exposto ao segundo, então o segundo bairro evocaria um juízo que pertence ao tipo relevante.”¹¹⁰ (CARRUTHERS, 2019, p. 157).

Em outras palavras, o que garantiria o valor de verdade do juízo seriam as mesmas disposições relevantes de João envolvidas nos juízos acerca dos dois bairros visitados sob as mesmas condições.

Voltando à consciência, Carruthers expressa da seguinte forma a semântica dos juízos que atribuem estados mentais conscientes:

[...] [A]s condições de verdade de juízos como, “A criatura *C* possui estados perceptuais que são *este-E*” deve ser: [...] “Se as disposições para julgar que

¹¹⁰ “[...] if one were to visit that other neighborhood under the same conditions with the same precursors (thus creating the same set of disposition-to-judge) then one would (or would not) count that other neighborhood as the same in the relevant respect. Or, somewhat more abstractly: if someone with the exact same dispositions-to-judge that were involved in one’s judgment about the first neighborhood were to be exposed to the second, then the second neighborhood would evoke a judgment that belongs to the relevant sort.”

subjazem meu uso do conceito ESTE-E fossem instanciadas na criatura C, então elas emitiriam um juízo de que alguns dos estados perceptuais da criatura são *este-E*.”¹¹¹ (CARRUTHERS, 2019, p. 157)

Ou, em outros termos:

[...] [A] condição de verdade para o pensamento, “O agente A subjaz estados *deste* tipo” (onde *DESTE* é um conceito fenomenal) é a seguinte: “Se as disposições subjacentes ao meu uso atual do conceito *DESTE* foram instanciadas na mente do agente A, então essas disposições emitiriam um juízo positivo de que *deste* está presente.”¹¹² (CARRUTHERS, 2020a, p. 12)

Quando se trata de juízos atribuindo consciência a outros humanos, fica claro que o contrafactual é satisfeito. Ora, é bastante razoável supor que dadas certas condições iguais ou similares, os outros humanos possuiriam as mesmas disposições para julgar que as minha quando utilizo um conceito fenomenal. Por exemplo, quando vejo o azul da pintura de Rothko, posso utilizar um indicador que se refere a minha experiência prazerosa de perceber a “*azulidade*” da pintura. O emprego do indicador também pressupõe o uso de capacidades como a da reflexão sobre o que penso da pintura e a linguagem em que a expresse. Diante deste contexto, é razoável considerar que um outro humano também é capaz de realizar uma experiência de cor similar à referida pelo meu indicador (claro, aqui excluo os indivíduos que possuem alguma deficiência em seu sistema visual, como os cegos e os daltônicos). Isso seria evidenciado pelas semelhanças que compartilhamos entre as nossas neurofisiologias, os nossos comportamentos, a nossa ontogenia cerebral e os nossos estados psicológicos. No entanto, não só os fatores em terceira-pessoa serviriam como evidência, mas também a capacidade das outras pessoas conceberem zumbis e se sentirem desafiados com o problema difícil e a lacuna explanatória. Nesse sentido, afirma-se que *as mesmas disposições de julgar subjacentes aos meus conceitos fenomenais podem ser instanciadas na mente de outros humanos*. Ou seja, todas as disposições que subjazem o uso do conceito fenomenal no meu juízo “então *isto* é como experienciar *x*” também podem ser instanciadas na mente de outros indivíduos coespecíficos.

¹¹¹ “[...] *the truth-condition of a judgment like “Creature C has perceptual states that are **this-E**” might be this: [...] “If the dispositions-to-judge that underlie my use of the concept **THIS-E** were to be instantiated in creature C, then they would issue in a judgment that some of the creature’s perceptual states are **this-E**.”*

¹¹² “[...] *the truth condition for the thought, “Agent A undergoes states of **this** sort” (where **THIS** is a phenomenal concept) is as follow: “If the dispositions underlying my current use of the concept **THIS** were instantiated in the mind of agent A, then those dispositions would issue in a positive judgment that **this** is present.”*

Por exemplo, assumindo que João seja um humano neurotípico, o juízo “João possui estados afetivos do mesmo tipo que *este-R*” possui valor de verdade. O juízo é verdadeiro caso João de fato possua estados afetivos iguais ao que o indicador se refere e falso caso João não os possua.

Contudo, quando se trata de animais o contrafactual não se encontra mais disponível, pois dizer que um animal possui as mesmas disposições para julgar que tenho quando utilizo “*este-R*”, seria supor uma mente totalmente diferente da que os animais de fato possuem. Considere o seguinte exemplo de Carruthers (2019, p. 158). Suponha que eu pense que os macacos possuem estados perceptuais tais como *este-R*. Nesse caso, estou realizando um juízo acerca da mente do animal em questão, um macaco. De acordo com o condicional contrafactual, as condições de verdade que fundamentam tal juízo deve ser algo como: se as disposições para julgar que subjazem o meu uso do conceito fenomenal “*este-R*” fossem instanciadas na mente deste macaco, então estas disposições seriam evocadas pelos estados mentais do próprio macaco (CARRUTHERS, 2019, p. 158). Como o próprio Carruthers aponta, “[...] o antecedente deste condicional implica que não estou mais pensando acerca da mente do macaco, mas na mente de um outro tipo de criatura com capacidades cognitivas como as dos humanos.”¹¹³ (CARRUTHERS, 2019, p. 158). Ora, ao projetar o meu conceito fenomenal utilizado em primeira-pessoa na mente dos outros animais, estaria supondo que eles possuem uma mente como a minha. No entanto, é evidente que todos ou a maior parte dos animais em nosso planeta *não* possui uma mente como a minha. Tendo isso em vista, o contrafactual que fundamenta os julgamentos de atribuição de consciência se torna inválido. A pergunta sobre se os macacos possuem estados conscientes como *este-R* conteria uma falsa pressuposição, a saber, a de que os macacos possuem uma mente que não é a que de fato possuem. A atribuição de consciência aos outros animais se torna, portanto, semanticamente indeterminado. Esses juízos não seriam nem verdadeiro e nem falso, por razões semelhantes à quando se pergunta “você parou de fazer dieta?” a alguém que nunca fez qualquer tipo de dieta. Assim, novamente Carruthers conclui que não há questão de fato acerca da consciência em animais.

2.8 **Objecções a Carruthers**

No que se segue, fornecerei três argumentos contrários à posição de Carruthers. O primeiro consistirá em um experimento de pensamento que apontará contradições internas à argumentação de Carruthers. Os outros dois argumentos, por sua vez, demonstrarão que

¹¹³ “[...] *the antecedent of this conditional entails that I am no longer thinking about a monkey’s mind, but the mind of another sort of creature altogether, with cognitive capacities like those of humans.*”

Carruthers se compromete apressadamente com teses que também produzem contradições em sua posição, a saber, uma *semântica intencional não naturalizada*, que é inadequada ao seu projeto completamente naturalista da consciência, e a *suposta natureza tudo-ou-nada* que a consciência possuiria.

2.8.1 O Argumento do Cientista Alienígena

Suponha que em um certo dia uma nave alienígena apareça repentinamente. O veículo aterrissa e de dentro dele saem alguns indivíduos, nativos de um planeta oriundo de uma galáxia distante da nossa. Segundo o que os alienígenas nos comunicam, o motivo pelo qual eles vieram visitar a Terra é apenas trocar conhecimento de modo pacífico com os humanos. Um destes alienígenas que integram a missão se chama Ceter Parruthers. Em seu planeta natal, Parruthers é considerado o maior especialista das áreas correspondentes às nossas ciências cognitivas e à filosofia. O conhecimento de Parruthers acerca dessas áreas é extremamente superior ao que atualmente possuímos. Por exemplo, Parruthers conhece a solução de todas as questões que nos aflige sobre a mente, a intencionalidade e a consciência, como: o problema difícil, a lacuna explanatória, se a mente é modular, a constituição metafísica da representação mental, se a ação é parte constitutiva da percepção etc.

No entanto, existe uma diferença fundamental entre os *Homo Sapiens* e a espécie a que Parruthers pertence, a saber, a arquitetura cognitiva que cada uma possui. A espécie alienígena possui uma arquitetura cognitiva semelhante à nossa, porém mais complexa. Além dos alienígenas possuírem todos os sistemas consumidores que a nossa arquitetura cognitiva dispõe, eles as possuem em uma versão mais complexa e avançada. Por exemplo, Parruthers também possui os sistemas responsáveis pela tomada de decisão e o relato verbal, no entanto os seus níveis de complexidade são incomparáveis com as dos humanos. Ademais, Parruthers também detém sistemas consumidores que não possuímos, como, por exemplo, a capacidade de se comunicar telepaticamente (assumindo que esta capacidade não infrinja nenhuma das leis da natureza) e a percepção do mundo na banda do infravermelho. Assim, os alienígenas possuiriam a conjunção $P \wedge Q$, onde P seriam os nossos sistemas consumidores, porém mais complexos, e Q os sistemas consumidores adicionais que não possuímos. Isto posto, as perguntas que se colocam são as seguintes: Ceter Parruthers e seus coespecíficos possuem consciência? Ceter Parruthers e seus coespecíficos conseguiriam emitir juízos que atribuem consciência aos humanos e aos outros animais?

Lembre-se que segundo Carruthers (2019), quando Parruthers se pergunta sobre se humanos possuem estados como *este-R*, esse juízo de atribuição deve satisfazer o seguinte condicional: se as disposições para julgar que subjazem o meu uso do conceito fenomenal “*este-R*” fossem instanciadas na mente do humano, então estas disposições seriam evocadas pelos estados do próprio humano. De acordo com Carruthers (2019, p. 184), haveria questão de fato acerca da atribuição de consciência dos alienígenas a nós, *pois compartilhamos o mesmo núcleo relevante de arquitetura cognitiva*. Ou seja, embora a arquitetura alienígena e humana possua diferenças, estas seriam apenas de grau. Deste modo, seria possível para Parruthers projetar o seu conceito fenomenal utilizado em primeira-pessoa “*este-R*” em nossa mente (diferentemente do que ocorreria entre os humanos e os outros animais). Além disso, o inverso também valeria, na medida em que poderíamos projetar o nosso conceito fenomenal “*este-R*” na mente de Parruthers e seus coespecíficos.

Aqui acredito que Carruthers se equivoca. O motivo, em linhas gerais, é o seguinte. Como Carruthers (2019, pp. 151-152) expressa, quando se trata dos humanos com outros animais o contrafactual não valeria nem mesmo para os que são filogeneticamente próximos de nós, como os chimpanzés (*Pan troglodytes*) e bonobos (*Pan paniscus*), espécies que compartilham cerca de 98% dos seus genomas conosco (PRÜFER et al, 2012), pois apesar de possuímos uma arquitetura cognitiva relativamente similar, a nossa seria muito mais complexa. Desta forma, segundo Carruthers, quando comparados os estados mentais dos chimpanzés e bonobos com os nossos, a maior parte dos estados mentais desses organismos estariam mais próximos dos nossos estados mentais inconscientes. Como o próprio Carruthers afirma, ou esses animais seriam considerados destituídos de consciência ou:

[...] não haveria questão de fato sobre se eles são ou não fenomenalmente conscientes (porque seus estados são mais semelhantes ao compartilhamento global dadas algumas maneiras de estipular os pesos no espaço de semelhança, porém são mais semelhantes aos estados *inconscientes* humanos dadas outras, e, então, deixam de ser classificados como consciente sob todas as ponderações razoáveis das dimensões).¹¹⁴ (CARRUTHERS, 2019, p. 152)

¹¹⁴ “[...] *there will be no fact of the matter whether or not they are phenomenally conscious (because their states are more similar to global broadcasting given some ways of stipulating the weights in the similarity space, but are more similar to human **un**conscious states given others, and thus fail to be classified as conscious under all reasonable weightings of the dimensions).*”

Poder-se-ia dizer que, da mesma forma que não há fato a ser estabelecido se chimpanzés e bonobos possuem ou não consciência, a despeito de compartilharem uma arquitetura cognitiva similar a nossa, também não haveria fato a ser estabelecido sobre a eventual consciência alienígena, uma vez que possuir uma arquitetura cognitiva similar não seria suficiente. Com efeito, a relação entre a mente dos alienígenas e dos humanos seria ainda mais controversa, na medida em que: a) a diferença entre a complexidade da mente alienígena e a dos humanos seria muito superior do que a relação entre a complexidade da mente humana e a dos outros homínídeos e b) os chimpanzés e bonobos são filogeneticamente próximos de nós, ao passo que não compartilhamos laços filogenéticos com os alienígenas. Deste modo, mesmo que os alienígenas compartilhem uma arquitetura cognitiva relativamente semelhante à humana, devido à complexidade superior da arquitetura alienígena, os nossos estados mentais provavelmente estariam mais próximos dos seus estados mentais inconscientes. Os alienígenas enfrentariam a mesma questão imposta entre nós e os chimpanzés e bonobos, isto é, ou o *Homo Sapiens* consistiria em uma espécie destituída de consciência ou não haveria qualquer fato sobre se humanos são ou não conscientes.

No entanto, a relação inversa também poderia ser considerada. Ora, na medida em que seria muito difícil imaginar a projeção de nossas disposições que subjazem o uso em primeira-pessoa de “*este-R*” em uma espécie cuja arquitetura cognitiva é muito mais complexa, as próprias disposições que subjazem o uso dos alienígenas de “*este-R*” poderiam ser de tal forma diferentes a ponto de não ser possível se traçar um paralelo com as nossas próprias disposições do uso de “*este-R*”.

Considere agora um cenário semelhante ao anterior, mas com a diferença de que Ceter Parruthers, o genial cientista cognitivo e filósofo alienígena, possui uma arquitetura cognitiva completamente diferente da nossa. Como o próprio Carruthers (2019, p. 185) sugere, a diferença entre a nossa mente e a desse alienígena seria ainda mais radical do que entre a nossa e a de um cefalópode. Diante desse caso, Carruthers (2019, p. 185) afirma que o contrafactual que fixa as condições de verdade dos juízos que atribuem consciência a Parruthers também seria inválido. Assim como no caso dos animais, aqui haveria uma indeterminação semântica e, portanto, não existiria questão de fato se Parruthers possuiria ou não consciência. A razão para Carruthers é a seguinte: supor que Parruthers instancia as mesmas disposições para julgar que subjazem o meu uso do conceito fenomenal “*este-R*”, seria considerar que Parruthers possui uma mente diferente da que ele realmente possui. Em outras palavras, se as disposições para julgar que subjazem o meu uso do conceito fenomenal “*este-R*” fossem

instanciadas na mente deste alienígena, então estas disposições seriam evocadas pelos estados do próprio alienígena (CARRUTHERS, 2019, p. 158). Contudo, o antecedente do condicional implica que não estaríamos mais tratando da mente do alienígena, mas de um ser com a arquitetura cognitiva como a nossa. Sendo assim, não há como projetarmos o nosso conceito fenomenal na mente de Carruthers e seus coespecíficos sem que com isso se altere a mente deles para algo que ela de fato não é.

Aqui me parece que o contrafactual tem um resultado contraintuitivo. Embora Carruthers e os seus coespecíficos possuam a conjunção $P \wedge Q$, sendo P o conhecimento de todas as verdades relevantes sobre a mente, a intencionalidade e a consciência, e Q as características cognitivamente complexas como, comportamentos flexíveis, inteligência e capacidade superior de resolução de problemas, dentre outras, de acordo com o contrafactual de Carruthers não haveria questão de fato sobre a presença da consciência entre eles. Apesar de tender a concordar que neste caso o contrafactual se encontra indisponível, creio que expõe dois problemas da posição de Carruthers. Primeiro, ele acaba demonstrando uma certa limitação da teoria semântica que a argumentação de Carruthers se assenta. Isto é, o que motiva o resultado cético não são os fatos da realidade e as evidências relevantes disponíveis, mas a teoria semântica viciada que Carruthers utiliza nos juízos que atribuem consciência. E segundo, o contrafactual apresenta uma forma de *excepcionalismo semântico antropocentrista* que é implausível. Ou seja, o contrafactual é semanticamente por princípio programado para *apenas* fixar os valores de verdade dos juízos que atribuem consciência a organismos com uma arquitetura cognitiva cujo núcleo é *quase idêntico* ao nosso. O ser humano é por princípio o centro irredutível da semântica de Carruthers. Assim, a indeterminação é devido aos problemas internos da semântica que Carruthers utiliza, e não em razão de questões factuais envolvendo os outros animais. Isso se torna claro quando a semântica falha em estabelecer condições de verdade para juízos que dizem respeito a organismos que são filogeneticamente próximos de nós e que compartilham conosco uma abundância de características relevantes, como no caso dos chimpanzés e bonobos. Isso também se aplica ao caso dos seres com capacidades e conhecimentos superiores aos nossos, tal como os do alienígena Carruthers. Ademais, creio que mesmo alguns casos de humanos neurodivergentes cairiam na indeterminação semântica de Carruthers, uma vez que possuem arquiteturas cognitivas diferentes. Ora, o que aqui deveria se revisar é a semântica subjacente aos juízos supracitados, pois é ela que os torna viciosamente indeterminado.

Como ilustração, considere, *mutatis mutandis*, a seguinte analogia. Em uma certa sala existem inúmeras caixas de diferentes formatos, sendo algumas tão similares entre si que são indiscerníveis à percepção humana. Além das caixas, há também uma chave. Ao entrar na sala, o humano projetor dos objetos faz o seguinte jogo semântico ao visitante: “Para que a abertura dessas caixas seja considerada uma *candidata* à fechadura, ela deve ser capaz de receber a introdução da chave. Caso a chave abra a caixa, então ela é uma fechadura. Porém, se a chave não a abrir, logo a abertura não é uma fechadura.” No que diz respeito aos juízos que atribuem a propriedade de “ser uma fechadura” às aberturas das caixas teríamos os seguintes cenários: a) para as caixas onde a chave foi introduzida e conseguiu abri-la o juízo será verdadeiro; b) para as caixas onde a chave foi introduzida, mas não conseguiu abri-la o juízo será falso; c) para as caixas onde a chave não conseguiu ser introduzida os juízos não possuirão valor de verdade (uma vez que para ser *candidata* à fechadura, a condição da chave ser introduzida na abertura da caixa deve ser satisfeita). Deste modo, os juízos que atribuem a propriedade de “ser fechadura” somente teriam valor de verdade para as caixas cuja chave conseguir ser introduzida.

Um tempo depois, o visitante descobre que a chave foi projetada para somente ser introduzida a uma determinada abertura de uma caixa especial. Ele também descobre que as aberturas de certas caixas também são aptas a receberem chaves, porém distintas da chave original. Por outro lado, as aberturas de outras caixas não são aptas a receberem chave alguma. Assim, no caso das aberturas que *não* foram capazes de receber a chave original, mas que receberam outras chaves distintas, o motivo pelo qual os juízos que atribuíam a propriedade “ser fechadura” serem indeterminados foi devido à condição de satisfação que era *por princípio* viciada em sua formulação original, uma vez que ela foi feita para *exclusivamente* atender uma caixa especial. Ora, tanto no caso da fechadura, quanto no da consciência, os juízos que atribuem tais propriedades estão imbuídos de uma semântica que por ser viciada expõe a sua limitação.

2.8.2 A Semântica Intencional Não Naturalizada

A segunda objeção que gostaria de abordar diz respeito à *semântica intencional* que Carruthers se compromete. Como foi visto em parágrafos anteriores, Carruthers (2019, pp. 147-148) sustenta que aquilo o que fixa a referência de um determinado conceito fenomenal C_n é a *intenção do sujeito possuidor de C_n* . Por exemplo, quando penso “esta é uma experiência prazerosa”, a extensão do conceito fenomenal indicador baseado por contato “esta” é fixada por

aquilo o que tenho em mente, isto é, pela intenção deliberada que tenho quando utilizo “esta”. De acordo com Carruthers, quando utilizamos um conceito fenomenal temos apenas a intenção de nos referirmos “[...] às qualidades de que temos ciência em nós mesmos.”¹¹⁵ (CARRUTHERS, 2019, p. 148). Essas qualidades que experienciamos seriam em sua concepção os conteúdos não conceituais globalmente compartilhados. Em resumo, é em razão de *intencionalmente* nos referirmos a apenas essas qualidades que os conceitos fenomenais não poderiam se referir a outros fenômenos como, por exemplo, tipos naturais (tal como “água” se refere a H₂O).¹¹⁶

No entanto, Birch (2020b) aponta que esse compromisso que Carruthers realiza suscita um dilema diante do arcabouço teórico em que ele está inserido. Por Carruthers ser um irrealista de *qualia*, a fixação da referência dos conceitos fenomenais por meio da intenção do sujeito possuidor do conceito iria de encontro a aquilo o que ele defende. Ora, como Birch (2020b) afirma, quando utilizamos um conceito fenomenal não temos em mente conteúdos não conceituais amplamente compartilhados em um espaço de trabalho global, mas às nossas experiências subjetivas, independentemente das suas naturezas finamente especificadas. Assumindo a tese nativista dos princípios dualistas, poder-se-ia até mesmo afirmar que temos em mente propriedades não físicas, como os *qualia* (BERENT, 2020; BLOOM, 2004). Aqui se encontra o dilema que Carruthers deve enfrentar. Se são as nossas intenções quando utilizamos um conceito fenomenal que fixam a sua referência: 1) ou o irrealismo de *qualia* está errado, uma vez que quando utilizamos um conceito fenomenal temos a intenção de nos referirmos às nossas experiências subjetivas, como os *qualia*, e não aos conteúdos não conceituais globalmente compartilhados; 2) ou, de fato, as nossas intenções *não* fixam o seu referente. Segundo a objeção de Birch (2020b), na medida em que o irrealismo de *qualia* e a semântica intencional são incompatíveis entre si, Carruthers deve endossar uma ou outra, não podendo assumir ambas ao mesmo tempo. Colocando em outros termos, ao endossar o irrealismo de *qualia* e rejeitar a semântica intencional, Carruthers abriria espaço para que os conceitos fenomenais pudessem se referir a, por exemplo, tipos naturais, como Bayne e Shea (2020)

¹¹⁵ “[...] *to the qualities we are aware of in ourselves.*”

¹¹⁶ A semântica de Carruthers guarda semelhanças com o programa de Grice (1957) sobre significado. Para Grice, grosso modo, o significado *P* de um enunciado *Q* é fundamentado na intenção *M* do proferidor *R* ao enunciar *Q* a uma audiência *S*. No entanto, existem diferenças importantes entre as duas propostas. Primeiro, Carruthers utiliza a semântica a fim de *fixar a referência* do conceito. Segundo, Carruthers não está tratando de enunciados, mas de um objeto peculiar que é o conceito fenomenal. Terceiro, enquanto para Grice (1957, pp. 386-388) a intenção não precisa necessariamente ser realizada conscientemente, para Carruthers (2019, p. 123) a articulação de conceitos fenomenais *só* pode ser feita explicitamente. Assim, dado que a articulação dos conceitos fenomenais só pode ser realizada de forma explícita, então a intenção de uso deles também deve ser realizada de forma explícita. Não analisarei a teoria do significado de Grice, pois fugiria do escopo da seção.

sustentam. Por outro lado, caso Carruthers endossasse a semântica intencional e rejeitasse o irrealismo de *qualia*, ele abriria espaço para que a consciência pudesse ser constituída por propriedades especiais, como os *qualia*. Se isso fosse o caso, a GWT não consistiria na teoria completamente reducionista e naturalizada que Carruthers almeja.

Acredito haver ainda um problema adicional. Ao longo de boa parte dos seus textos recentes sobre o tema, mais notadamente o livro de 2019 “*Human and Animal Minds: The Consciousness Questions Laid to Rest*”, Carruthers busca estabelecer um empreendimento filosófico *naturalista robusto*, levando a sério os dados empíricos disponíveis. Contudo, em relação à semântica adotada, creio que Carruthers não se atenta ao fato de apelar acriticamente a apenas, ou em larga medida, as suas intuições privadas sobre o que as pessoas pensam quando utilizam um conceito fenomenal (assumindo que existam conceitos fenomenais). Não me parece tão claro que ao utilizarmos um conceito fenomenal apenas queremos nos referir às qualidades da experiência, sendo elas conteúdos não conceituais globalmente compartilhados. Ou mesmo se queremos nos referir aos *qualia*, como Birch (2020b) coloca. Ao utilizar um conceito fenomenal, um indivíduo pode querer *intencionalmente* se referir a coisas completamente distintas e incompatíveis com aquilo ao que os seus coespecíficos querem se referir ao utilizarem um conceito fenomenal igual ou similar. Colocando de forma mais geral, em uma ocasião *O*, um indivíduo *S* ao utilizar um conceito fenomenal *C* pode ter em mente um referente *R* que, por sua vez, é completamente distinto do referente *P* a qual o indivíduo *Q* tem em mente ao utilizar um conceito fenomenal *C'* que é similar ou igual a *C*, em uma ocasião *O'* que é igual ou similar a *O*. Por exemplo, ao comer um sorvete na praia o seguinte pensamento pode ser gerado: “*esta* é uma experiência prazerosa”. Diante desse pensamento, um certo pampsiquista poderia argumentar o seguinte: ao utilizar o conceito fenomenal “*esta*”, ele apenas tem por intenção se referir a certas propriedades físicas, porém intrínsecas da realidade que não são captáveis pelas ciências. Um dualista de substâncias, por sua vez, poderia argumentar que deseja se referir a propriedades fenomenais que não são físicas. Um proponente da cognição corporificada poderia expressar a intenção de somente se referir ao corpo como um todo. Mas e quanto aos chamados “leigos”? Carruthers poderia argumentar que ao utilizar o termo “*nós*” ele apenas busca se referir à massa de pessoas que não são especialistas em filosofia e/ou nas ciências da mente. Todos os especialistas anteriormente citados não fariam, portanto, parte do seu recorte, em virtude de estarem viesados pelas teorias a que defendem. Isso tampouco me convence. Caso fosse feita uma afirmação desse tipo, ela me pareceria completamente *ad hoc*, dado que o único intuito dela ser proferida é a satisfação da semântica que embasa os seus

argumentos. Ademais, nada garante que a massa de pessoas leigas possuiria as mesmas intenções e disposições que Carruthers ao utilizar um conceito fenomenal para se referir a *P* e não *Q*.

Dado o espírito naturalista do empreendimento de Carruthers¹¹⁷, o mais correto e parcimonioso seria, ao que me parece, fundamentar a semântica em pesquisas empíricas sobre o que de fato as pessoas têm em mente quando utilizam um conceito fenomenal. Ainda assim alguns problemas permaneceriam em aberto. Para que o resultado dessa pesquisa empírica pudesse ser considerado minimamente razoável, ela deveria seguir os procedimentos mais rigorosos que as áreas concernentes ao tema demandam¹¹⁸, como, por exemplo, ter o *n* amostral suficientemente grande e qualitativamente relevante, abarcando não apenas populações com psicologia consideradas *WEIRD*'s (acrônimo em inglês utilizado para se referir aos indivíduos que cresceram em sociedades do tipo ocidental, escolarizada, industrializada, rica e democrática) (HENRICH, 2021), mas também as populações de sociedades não *WEIRD*, como as denominadas “tradicionais” (e. g. as sociedades de caçadores-coletores). Ainda que a pesquisa fosse realizada com sucesso e indicasse que ao utilizar um conceito fenomenal a maioria das pessoas têm em mente as tais qualidades a que Carruthers se refere, ainda assim teríamos as que afirmam possuírem outra coisa em mente quando utilizam um mesmo conceito fenomenal. Poderíamos ter até mesmo pessoas dizendo que quando utilizam um conceito fenomenal querem se referir ao mesmo tempo a qualidades experienciais e a tipos naturais, como Bayne e Shea (2020) argumentam ser possível. Não me parece claro como Carruthers poderia conciliar as suas intuições apriorísticas e privadas com aquilo o que de fato as pessoas pensam, ou acham que pensam.¹¹⁹

Ademais, Carruthers (2019) não justifica o uso dessa semântica. Ora, por que deveríamos aceitar essa e não uma outra semântica que seja empiricamente motivada, dado o projeto naturalista que o próprio Carruthers advoga, e melhor lide com o problema que estamos enfrentando? A indeterminação semântica poderia ser, na verdade, um problema da semântica

¹¹⁷ Devo ressaltar que Carruthers é inegavelmente um dos filósofos contemporâneos com o maior domínio da literatura empírica sobre a mente. Não é uma mera contingência que os seus trabalhos sejam amplamente citados e utilizados por cientistas de áreas como as da psicologia, neurociência e ciências cognitivas. Por essa e outras razões que muitos o caracterizam não apenas como um “filósofo” *stricto sensu*, mas também um “cientista cognitivo”. Portanto, não é o meu intuito afirmar que ele não possua conhecimento sobre as pesquisas sobre o tópico, mas que apenas o ignora quando diante do debate sobre a semântica intencional que levanto conjuntamente com Birch (2020b).

¹¹⁸ Incluo aqui sobretudo a psicologia, a antropologia e a filosofia experimental.

¹¹⁹ Poder-se-ia questionar até mesmo a validade desses tipos de experimentos que de alguma forma recorrem à introspecção, na medida em que ela é falível e altamente passível de cometer erros (NISBETT e WILSON, 1977). O próprio Carruthers (2011) é extremamente cético quanto a nossa capacidade de introspecção, uma vez que, segundo ele, cometemos erros sistemáticos acerca dos nossos próprios pensamentos.

que fundamenta as condições de verdade dos juízos que atribuem consciência, e não a uma suposta característica *sui generis* que o tópico da consciência animal detém. Ao fim e ao cabo, a conclusão radical de Carruthers poderia ser nada mais que a limitação de sua semântica sendo exposta.

Isto posto, creio que o dilema de Carruthers permanece em aberto. Ou de fato as nossas intenções não fixam o referente dos conceitos fenomenais ou fixam, mas não sabemos quais são as intenções que corretamente o fazem. Dizer que as intenções que corretamente fixam os referentes são aquelas que tem como direção os conteúdos não conceituais globalmente compartilhados seria cometer uma petição de princípio.

2.8.3 Seria a Consciência um Fenômeno Tudo-ou-Nada?

Ao longo de sua argumentação, Carruthers defende que a consciência é um fenômeno do tipo tudo-ou-nada e que, portanto, vai ao encontro daquilo o que a GWT sustentaria. Ora, isso não me parece óbvio. No que se segue, argumentarei que existe uma incompatibilidade entre a versão reducionista da GWT que Carruthers propõe e a tese da natureza tudo-ou-nada da consciência. Os meus argumentos são os seguintes:

Argumento 1:

- A) A consciência é um fenômeno físico tudo-ou-nada ou gradual.
 - B) Se a consciência é um fenômeno físico tudo-ou-nada, então ela não pode ser um fenômeno físico que admite graus.
 - C) A consciência é um fenômeno físico tudo-ou-nada.
-

D) Logo, a consciência não é um fenômeno físico gradual.

Argumento 2:

- A) A consciência consiste apenas em conteúdos representacionais compartilhados globalmente.
 - B) Os conteúdos representacionais são fenômenos físicos que admitem graus.
-

C) Portanto, a consciência é um fenômeno físico gradual.

D) A partir dos argumentos 1 e 2 conclui-se que a consciência é um fenômeno físico tudo-ou-nada e gradual ao mesmo tempo.

Vejam os o que cada premissa do primeiro argumento aborda. A premissa (A) expressa uma instanciação do Princípio do Terceiro Excluído da Lógica Clássica ($\varphi \vee \neg\varphi$). A premissa pode ser formalizada da seguinte forma. Considerando P: consciência; Q: um fenômeno físico tudo-ou-nada; $\neg Q$: um fenômeno físico gradual; têm-se que:

$$\forall x (Px \rightarrow ((Px \wedge Qx) \vee (Px \wedge \neg Qx)))$$

Lê-se a fórmula acima como: para todo x ,¹²⁰ se x for consciência, então ou x é consciência e é um fenômeno físico tudo-ou-nada ou x é consciência e um fenômeno físico gradual. Ou seja, a consciência é um fenômeno tudo-ou-nada ou a sua negação, que é ser um fenômeno gradual. A premissa (B) enuncia que se a consciência é, primeiro um fenômeno físico e de natureza tudo-ou-nada, tal como Carruthers sustenta, por conseguinte a consciência não pode ser a negação deste último tipo de fenômeno físico, isto é, ser um fenômeno físico e gradual. Algumas considerações devem ser feitas. A primeira é de que a premissa considera que a consciência é um fenômeno físico e que a sua negação também o é. Por sua negação entendo não a possibilidade de a consciência ser um fenômeno não físico, mas ser um fenômeno físico cuja natureza é diferente daquela que foi anteriormente considerada. Ora, considerar que a negação da consciência ser um fenômeno físico do tipo tudo-ou-nada seja ser um fenômeno não físico tudo-ou-nada, ou ser um fenômeno não físico gradual, tornaria o argumento menos plausível e interessante, uma vez que Carruthers assume que a consciência é física. Portanto, considero que a negação de ser um fenômeno físico tudo-ou-nada é ser um fenômeno físico que admite gradações. Retomando, na premissa (B) é admitido que existiriam dois tipos de fenômenos físicos: os de natureza tudo-ou-nada e os graduais. Alguns autores, notadamente Tye (2021), argumentam que quase todos os fenômenos físicos são graduais, com exceção daqueles que pertencem à microrrealidade (reino físico mais fundamental que o quântico, onde estariam dispostas as quiddidades, propriedades intrínsecas, inefáveis e especiais que por princípio não seriam captáveis pelas ciências). No entanto, para fins da presente argumentação, deixo em suspenso a existência desse reino misterioso que Tye (2021, 2022) e outros pampsiquistas postulam (CHALMERS, 2017; GOFF, 2017). Nesta premissa o que realmente

¹²⁰ Em universo de discurso U tal que $\forall x (x \in U \rightarrow Px)$.

importa é a existência de fenômenos físicos que são tudo-ou-nada e outros que são graduais e que se a consciência for um fenômeno físico do tipo tudo-ou-nada, ela não poderá ser um fenômeno físico gradual. A premissa (C) expressa apenas a posição de Carruthers segundo a qual a natureza da consciência é física e tudo-ou-nada. Por fim, a conclusão (D) se segue das premissas, uma vez que se a consciência é um fenômeno físico tudo-ou-nada ela não pode ser a sua negação (ser um fenômeno físico gradual). Observe que o argumento 1 não seria controverso à luz do que Carruthers defende, pois ele nada mais é que o resumo de proposições que o próprio autor aceita.

Vejamos agora o que as premissas do segundo argumento tratam. A premissa (A) expressa o compromisso de Carruthers com o irrealismo de *qualia* e à GWT. Como foi visto anteriormente, para Carruthers (2019, pp. 116-120) a consciência fenomenal deve ser explicada de forma *completamente reducionista* em termos *representacionais*. A consciência é reduzida a conteúdo representacional não conceitual globalmente compartilhado, e não a certos estados neuronais e/ou mecanismos cerebrais. Em outras palavras, para Carruthers (2020a, p. 7) a consciência nada mais é que conteúdo representacional compartilhado globalmente aos sistemas consumidores disponíveis em um certo tipo de arquitetura cognitiva. Carruthers, portanto, rejeita a tese de que a consciência envolve propriedades especiais e endossa a tese de que a consciência pode ser totalmente reduzida a conteúdos representacionais que desempenham um certo papel causal. Note que Carruthers realiza uma interpretação cognitivista da GWT, na medida em que reduz a consciência a representações que desempenham um certo papel causal. Deste modo, ele se afasta das interpretações neuronais da GWT (DEHAENE, 2014; MASHOUR et al, 2020), e se aproxima da interpretação clássica de matriz cognitivista (BAARS, 1988). Ademais, é importante lembrar que Carruthers (2019, pp. 118-119) sustenta que o conteúdo representacional e as representações não precisam ser reduzidos a entidades de nível inferior, como as da biologia, da química ou da física. O conteúdo representacional e as representações seriam termos primitivos, isto é, irredutíveis a propriedades de nível inferior. As razões fornecidas para se comprometer com essa forma de primitivismo seriam basicamente as seguintes. O primeiro argumento afirma que: conteúdo representacional é um dos postulados mais básicos da psicologia cognitiva, as propriedades psicológicas são reais, o progresso de integração entre a psicologia e as neurociências também é real, logo seria plausível considerar que os conteúdos representacionais também sejam reais. O segundo argumento, por sua vez, expressa que os conteúdos representacionais e as representações já são consideradas propriedades “aceitavelmente naturalizadas”

(CARRUTHERS, 2017b, p. 75) que desempenham um papel central e com grande sucesso explanatório nas ciências cognitivas, psicologia e neurociência. Deste modo, Carruthers assume que conteúdos representacionais e representações são propriedades primitivas já naturalizadas, não precisando ser reduzidas a propriedades de nível inferior, como as da biologia, química e física.

A premissa (B) afirma que os conteúdos representacionais são fenômenos físicos que admitem graus. Considere dois exemplos retirados de Tye (2021, p. 7), porém levemente modificados. Considere a propriedade de representar “piscina”. Durante um determinado período histórico, a palavra “piscina” significava algo como um viveiro de peixes. Após um certo tempo, “piscina” passou a significar aquilo o que hoje intentamos quando utilizamos tal palavra, a saber, um certo reservatório de água utilizado para banho. Afirmar que a mudança de significado que a palavra sofreu foi abrupta, isto é, mudou de 0 para 1 me parece ser um tanto quanto implausível. O que provavelmente ocorreu foi que a alteração do significado se deu de forma gradual durante um determinado período de tempo e espacialmente distinta, onde em um ponto do tempo e do espaço geográfico não era claro se a palavra definitivamente significava uma coisa ou outra. Por exemplo, a mudança pode ter ocorrido primeiro no Rio de Janeiro e após um determinado tempo em Recife. Durante um certo momento existiam casos limítrofes onde era indeterminado se “piscina” representava viveiro de peixes ou reservatório de água para banho. O motivo geral, como bem aponta Tye (2021, p. 7), é o de que a própria propriedade de representar algo admite casos limítrofes. Em outras palavras, a propriedade de representar x é um fenômeno físico de natureza gradual. Para tornar mais claro, considere agora um determinado conjunto C de neurônios no córtex V1 disparando conforme representam um objeto O no campo de visão de um indivíduo neurotípico. Em uma concepção padrão de representação, diz-se que representar envolve rastreamento sob certas condições normais (TYE, 2021, p. 7). Ora, dado que o conceito utilizado de “normalidade” também é gradual, a própria noção de representação também será gradual. Seja se “normalidade” for entendido sob o prisma “[...] estatístico (o que *tipicamente* acontece)”¹²¹ (TYE, 2021, p. 7) ou “[...] teleológico (o que *supostamente* deveria acontecer caso o sistema relevante esteja operando propriamente)”¹²² (*ibidem*).

Ademais, como Souza Filho (no prelo) bem aponta, poder-se-ia argumentar que assim como ocorre com as categorias biológicas, os estados representacionais também surgem

¹²¹ “[...] statistically (what **typically** happens).”

¹²² “[...] teleologically (what is **supposed** to happen if the relevant system is operating properly).”

em graus. Considere um certo traço biológico T , como, por exemplo, uma asa e admita uma noção etiológica de função biológica, segundo a qual um traço biológico T possui a função F se e somente se ele foi selecionado, via seleção natural, para realizar um efeito E (NEANDER, 1991). De acordo com a teoria da evolução, T surge em graus, isto é, a evolução de T se dá através da acumulação de pequenas e diversas mutações genéticas ao longo de diversas gerações de populações de organismos O em um período de tempo Q (COYNE, 2014, pp. 23-25; DAWKINS, 2001; FUTUYMA, 2009, pp. 677-680). Nesse sentido, o traço biológico “asa” não surgiu em um único rompante, mas através de sucessivos acúmulos de mutações que se tornaram mais frequentes em uma certa população de indivíduos em virtude de terem sido favorecidas pela seleção natural (uma vez que contribuía de alguma forma ao sucesso reprodutivo e/ou de sobrevivência dessa população). Assim, o traço ancestral de “asa” provavelmente não possuía a função de voo, mas antes alguma outra função (e. g. planar, cortejo ou termorregulação). Após um certo tempo, mutações genéticas em sucessivas gerações fizeram com que surgisse o traço “asa” cuja função é fazer com que o seu portador voasse. Existiu em um certo momento da história evolutiva de T , ocorrências de traços que não eram completamente uma asa, mas também não eram algo muito diferente disso. Ou seja, existiriam casos limítrofes de asa, onde não seria claro se um indivíduo a possui ou não. Ora, se os estados e conteúdos representacionais também estão sujeitos ao ácido universal do darwinismo (DENNETT, 1995), como provavelmente é o caso, então os estados e conteúdos representacionais provavelmente surgiram através de sucessivas modificações ao longo de um período de tempo em diversas gerações e com uma função diferente da que possuem atualmente. Assim como ocorreu com os outros traços biológicos, também existiu uma zona cinzenta onde era indeterminado se certos estados mentais E eram ou não representacionais. Ou seja, haveria casos limítrofes de sistemas representacionais (SOUZA FILHO, no prelo).

No entanto, Carruthers poderia argumentar que assim como ocorre no caso da consciência, seria difícil de se conceber casos parcialmente representacionais ou limítrofes. Novamente o apelo a intuições relacionadas à concebilidade é frágil, pois assim como muitos autores argumentam que estados mentais conscientes são graduais e que é possível conceber casos limítrofes de consciência (GODFREY-SMITH, 2020c; SCHWITZGEBEL, 2021; TYE, 1996, 2021; VEIT, 2022a), o mesmo poderia ser levantado em relação aos estados e conteúdos representacionais. Ademais, o próprio fato de que os estados e conteúdos representacionais estão sujeitos à evolução forneceria razões a favor de uma visão gradualista acerca deles. Assumindo o fisicismo, não haveria nada de especial em estados e conteúdos representacionais

que o separem de todas as outras categorias físicas. Concluo que estados e conteúdos representacionais de fato admitem gradações e podem apresentar casos limítrofes.

A conclusão (C) se segue das premissas, posto que se a consciência consiste em nada mais que conteúdo representacional sendo globalmente compartilhado e a natureza de tal conteúdo admite gradações, logo a consciência é um fenômeno cuja natureza é gradual. No entanto, como vimos no Argumento 1, a consciência não pode ser um fenômeno gradual. Ora, o que afinal de contas a conclusão (D) revela é uma contradição, pois a consciência seria ao mesmo tempo um fenômeno tudo-ou-nada (Argumento 1) e gradualista (Argumento 2).¹²³

Isto posto, não vejo meios hábeis para que Carruthers se esquive com sucesso dessa contradição, na medida em que ele estabelece uma relação de *identidade* e *redução* entre a consciência e conteúdos representacionais. Assim, como meio de fuga, Carruthers não poderia apelar a outros tipos de relações entre a consciência e conteúdos representacionais, como, por exemplo, as de fundamentação (*grounding*) ou superveniência.¹²⁴ Por exemplo, ele não pode argumentar que a consciência é um fenômeno tudo-ou-nada que está metafisicamente fundamentado (*grounded*) em conteúdos representacionais que, por sua vez, admitem gradações. Onde dizer que:

[...] a propriedade P é metafisicamente fundamentada (*grounded*) em uma propriedade Q é dizer que (a) é metafisicamente necessário que toda vez que Q é instanciado, P também o é, e (b) que a natureza de P explica o porquê P é metafisicamente necessitado por Q.¹²⁵ (TYE, 2021, p. 9)

Os dois principais motivos são, como já foi visto neste capítulo, os de que Carruthers 1) estabelece relações de *identidade* e *redução* entre a consciência e conteúdos representacionais e 2) ele se compromete com a tese primitivista no que tange conteúdos representacionais e representações.¹²⁶ Os conteúdos representacionais e as representações não são redutíveis a propriedades de nível inferior que, por algum motivo, seriam fenômenos físicos

¹²³ Considerando P: consciência; Q: fenômeno físico tudo-ou-nada; \neg Q: fenômeno físico gradual, têm-se o seguinte: $\exists x (Px \wedge Qx \wedge \neg Qx)$. Onde existe um x tal que x é a consciência e também é um fenômeno físico tudo-ou-nada e a sua negação ao mesmo tempo, isto é, ser um fenômeno físico gradual.

¹²⁴ Para uma introdução à fundamentação (*grounding*), cf. McKenzie, 2022. Para a superveniência, cf. McLaughlin e Bennett, 2018.

¹²⁵ “[...] a property P is metaphysically grounded in a property Q is to say (a) that it is metaphysically necessary that whenever Q is instantiated, P is too, and (b) that the nature of P explains why P is metaphysically necessitated by Q.”

¹²⁶ Ademais, é altamente debatível se é possível estabelecer uma relação de fundamentação (*grounding*) entre a consciência, tal como um fenômeno tudo-ou-nada, e uma propriedade física x do tipo gradual. Cf. Tye, 2021.

tudo-ou-nada.¹²⁷ Portanto, caso Carruthers se comprometa com a tese tudo-ou-nada da consciência, ela não poderá ser reduzida a conteúdo representacional, o que resultaria no fracasso de seu projeto reducionista da consciência. Por consequência, seriam invalidados os seus argumentos de que não haveria questão de fato em relação à consciência nos animais. Ao fim e ao cabo, me parece que Carruthers quer comer o bolo ao mesmo tempo em que deseja que ele permaneça intacto, mas como já é sabido, isso é impossível de ser feito.

Concluo que além da posição cético-radical de Carruthers ser pouco razoável, ela admite inconsistências internas que, como busquei demonstrar, inviabilizam o seu projeto reducionista e naturalista da consciência. Considerando os argumentos fornecidos neste capítulo, creio ser justificável que se abandone essa posição. As pesquisas sobre consciência animal devem, portanto, continuar prosseguindo e avançando. Isto posto, existem outras posições, em especial as pertencentes à abordagem epistêmica, que são mais promissoras para o uso nas pesquisas sobre consciência animal. No próximo capítulo, examinarei algumas delas e defenderei aquela que considero ser atualmente a mais plausível.

¹²⁷ Alguns autores consideram que, com exceção das entidades pertencentes à microrrealidade, todos os fenômenos físicos admitem gradações (TYE, 2021, 2022). Se este for o caso, então qualquer tentativa por parte de Carruthers de identificar e reduzir a consciência a fenômenos físicos seria fadada ao fracasso, uma vez que todos os fenômenos físicos que lhe interessam admitem gradações.

3 A ESTRATÉGIA DOS MARCADORES DE CONSCIÊNCIA

Neste capítulo, tratarei especificamente da abordagem epistêmica. Em linhas gerais, a abordagem epistêmica sustenta que devemos buscar um método que seja capaz de justificar de forma abductiva a hipótese H sobre a presença da consciência em um dado animal como a melhor explicação disponível à luz das evidências. Para isso, ela se vale da chamada “estratégia dos marcadores de consciência”, um procedimento que utiliza indicadores supostamente associados à consciência que justificariam H . Como forma de clarificar a estratégia, buscarei apresentar brevemente no que consiste a inferência para a melhor explicação e os marcadores de consciência. Para evitar possíveis confusões, também distinguirei os marcadores de consciência dos chamados “correlatos neurais da consciência” e dos “constituintes da consciência”. Em seguida, examinarei duas posições epistêmicas relevantes: a abordagem padrão e a abordagem newtoniana de Tye (2017, 2018). Fornecidas as devidas razões para que se abandone essas duas propostas, defenderei a “abordagem teórica leve”, tese epistêmica que considero ser mais a plausível atualmente (BIRCH, 2020a). Por fim, buscarei responder a potenciais objeções contra ela.

3.1 Introdução

Como viu-se no 1º capítulo, quando se trata da localização da consciência na árvore filogenética e qual a melhor forma de indicarmos a presença da consciência entre os animais, respectivamente o problema da distribuição e da mensuração, existem ao menos três conjuntos de respostas possíveis: a abordagem teórica, a abordagem epistêmica e a abordagem da percepção direta. Em linhas gerais, a abordagem teórica propõe a aplicação de uma teoria prévia sobre a consciência humana (T_n) como forma de indicar se um determinado animal possui ou não consciência. Seriam considerados conscientes somente aqueles animais que cumprissem os mesmos requisitos que T_n exige dos seres humanos. Como Shevlin (2021) distingue, existem interpretações diferentes da abordagem teórica, onde umas são mais radicais ou moderadas que as outras. No capítulo anterior, examinei a versão cético-radical proposta por Peter Carruthers (2019). Grosso modo, Carruthers sustenta que não haveria questões de fato sobre se animais possuem ou não consciência. Segundo o autor, o projeto de investigar a consciência animal já iniciaria condenado ao fracasso. Demonstrei que a sua posição incorre em contradições e que, portanto, devemos abandoná-la.

A abordagem da percepção direta, por sua vez, sustenta que percebemos de maneira imediata os estados mentais conscientes dos outros animais. Ou seja, é apenas em virtude da

nossa interação com o animal e/ou a percepção de suas expressões que detectamos sem mediações de inferências em nossa mente os seus estados conscientes. Por fim, a abordagem epistêmica sustenta que não precisaríamos nos comprometer com T_n e tampouco com algum mecanismo inato, como um detector animista. A abordagem epistêmica se assenta na chamada “*estratégia dos marcadores de consciência*” (EMC). Segundo ela, são os fatores pré-teóricos que justificam abduktivamente se o animal em questão é ou não consciente. Esses fatores pré-teóricos são os chamados *marcadores de consciência*. É importante ressaltar que a utilização da EMC varia de acordo com a versão da abordagem epistêmica adotada. O presente capítulo será dedicado a esta abordagem.

Antes de prosseguir ao exame mais detalhado das diferentes formas da abordagem epistêmica, algumas distinções e clarificações serão feitas. No que se segue, apresentarei brevemente o que é a inferência para a melhor explicação, utilizada amplamente por essa abordagem. Logo após, distinguirei os conceitos de correlatos neurais, marcadores e constituintes da consciência, para que, enfim, adentremos no assunto alvo deste capítulo.

3.2 Distinções e Clarificações

3.2.1 Inferência para a Melhor Explicação

A inferência para a melhor explicação (IME), ou inferência abduktiva, é a forma mais comum de justificção de hipóteses que encontramos nas ciências.¹²⁸ Diferentemente dos argumentos dedutivos, a conclusão da IME possui um caráter ampliatiivo do conhecimento, de modo similar ao dos argumentos indutivos. Na IME, bem como no argumento indutivo, a sua conclusão não se segue necessariamente das premissas. Ela é feita com base nas evidências relevantes disponíveis sobre o assunto e nas chamadas “virtudes teóricas”, como a simplicidade, a coerência e a generalidade.¹²⁹

Considere o seguinte exemplo. Suponha que S seja um humano neurotípico em nosso mundo atual. Em uma dada madrugada, S acorda, levanta-se de sua cama e anda rumo à geladeira em busca de água. Ao abrir a geladeira, S pega uma garrafa de água e ingere o líquido contido no recipiente. Podemos concluir que a razão de S ter acordado na madrugada e ter ido em busca de água foi devido à sua sede. Essa é a melhor hipótese, dado que ela é a que melhor explica os dados disponíveis, mais se adequa às virtudes teóricas e é suficientemente boa. Note

¹²⁸ O intuito aqui é apenas apresentar brevemente o assunto. Para a obtenção de informações mais detalhadas, cf. Douven, 2021; Lipton, 2004; Tye, 2017, cap. 4.

¹²⁹ É debatível no que consistem as virtudes teóricas ou mesmo se elas existem. Para maiores detalhes, cf. Godfrey-Smith, 2021, cap. 13.

que a conclusão do argumento não estava contida nas premissas, ela foi feita de acordo com a IME. Como forma de explicar certos tipos de comportamentos, a psicologia popular assume como causas estados mentais como crenças e desejos. Isso pode ser observado através do exemplo supracitado, pois se *S* deseja *P* e crê que *C* leva a *P*, então, tudo o mais constante e sem outros desejos conflitantes, *S* faz *C*. Seria por meio da psicologia popular que prevejo com sucesso que *S* deseja água para satisfazer a sua sede, quando *S* pede um copo de água em um restaurante. Se não tivéssemos a psicologia popular, a capacidade de predição do comportamento dos outros indivíduos, e até de certos animais, seria muito dificultada, uma vez que não entenderíamos o funcionamento mental e como ele se relaciona com os comportamentos dos indivíduos. Suponha o seguinte caso. Ao chegar na reunião do condomínio às 18 horas, o síndico encontra todos os condôminos o esperando na sala 201. A explicação intuitiva dos condôminos terem previsto a hora exata que o síndico chegaria à reunião é de que o síndico *acredita* que a reunião se inicia às 18 horas na sala 201. Além disso, também se crê que o síndico possui os conceitos de “reunião”, “18 horas” e “sala 201”, e os consegue manipulá-los em seu pensamento. Assim, tudo isso contribuiria para gerar o comportamento do síndico ir à reunião do condomínio às 18 horas na sala 201. Como o filósofo Jerry Fodor já argumentou, a falta da psicologia popular em nossa vida resultaria na maior catástrofe intelectual de nossa espécie (FODOR, 1987, p. xii).

Claro, poder-se-ia argumentar que caso *S* estivesse sendo controlado por pequenos robôs malignos implantados em seu cérebro, *S* teria ido em busca de água, mesmo que não estivesse com sede. Contudo, o que mais estaria de acordo com tudo o que sabemos sobre o mundo atual, o funcionamento do corpo humano e a que apresenta as melhores virtudes teóricas é a hipótese de que *S* foi em busca de água porque estava com sede. A postulação de robôs malignos, ou quaisquer hipóteses semelhantes, não se coaduna com o que aqui denominei “virtudes teóricas”, como, por exemplo, o princípio da parcimônia explicativa. Na medida em que a melhor explicação possui uma economia de entidades e está mais de acordo com as evidências disponíveis.

No que toca o conhecimento acerca das outras mentes, os argumentos que utilizam a IME são muitas das vezes confundidos com os denominados “argumentos por analogia”, também conhecidos como “argumentos analógicos”. Vejamos o seguinte exemplo:

- A) Toda vez que sinto dor nos meus membros, me comporto de uma forma peculiar.

B) Observo que os outros indivíduos se comportam da mesma forma.

Deste modo, infiro por analogia que:

C) Provavelmente, as outras pessoas também possuem dor nos membros.¹³⁰

O argumento parte de um caso específico, os meus estados psicológicos e os meus comportamentos, para inferir que outras pessoas possuem estados mentais. O argumento por analogia levanta inúmeros problemas. O primeiro é a falácia da generalização. A conclusão de que outras pessoas também possuem estados mentais é inferida apenas do meu próprio caso, a saber, a minha dor nos membros. Com efeito, ninguém está autorizado a inferir que F 's são G 's com base apenas em um único caso em que um F é G (caso contrário, perceber um cisne negro na Austrália me autorizaria a supor que todos os cisnes são negros).

O segundo problema está relacionado com a segunda premissa. Apesar das outras pessoas serem em geral parecidas comigo, muitas das suas características são diferentes. Nem todo mundo apresentaria o mesmo comportamento que eu quando diante de um estímulo E_n , ou possuiriam as mesmas constituições neurológicas que as minha (pense em indivíduos com lesões cerebrais e outras condições neurológicas adversas). Novamente isso desautoriza a conclusão do argumento.

Em terceiro lugar, teríamos o problema da correção. A inferência indutiva segue de uma constatação da minha própria mente. No entanto, devido à sua própria característica peculiar, só podemos verificar os estados mentais que dizem respeito a nós mesmos. Não podemos, assim, verificar os estados mentais que não sejam os nossos próprios, isto é, não podemos checar independentemente os estados mentais de outros indivíduos (TYE, 2017, pp. 54-55). Um último problema, e talvez o mais fraco, que poderia ser levantado é a de como quantificar a probabilidade que a conclusão possui. O quão provável a conclusão tem que ser para que se torne razoável? E como quantificar a probabilidade de x , quando a variável diz respeito a um objeto tão peculiar como são os estados mentais?¹³¹

Em contrapartida, os argumentos do tipo IME podem até conter alguma analogia em uma de suas premissas, mas não é ela que desempenhará o papel inferencial. O que desempenhará o papel inferencial no argumento é a própria IME. Realizamos tal inferência

¹³⁰ Tye (2017, p. 54) apresenta um exemplo similar.

¹³¹ Uma resposta possível seria apelar ao bayesianismo. A vasta literatura da ciência da consciência possui métodos quantitativos interessantes. Contudo, não entrarei neste debate. Cf. Dehaene, 2014. Para uma introdução à relação entre bayesianismo e conhecimento científico, cf. Godfrey-Smith, 2021, cap. 12; Schupbach, 2022.

abdutiva com base no conhecimento disponível sobre o assunto, as evidências pertinentes disponíveis e o grau de confiança delas, incluindo nesse conjunto homologias entre os nossos estados mentais e os estados de terceiros em face de condutas semelhantes.

Alguns pesquisadores *interpretam* a IME por meio do Teorema de Bayes, formulado inicialmente pelo clérigo britânico Thomas Bayes e refinado posteriormente por Pierre-Simon Laplace. Outros autores *identificam* a própria IME com o Teorema de Bayes (GODFREY-SMITH, 2021; SETH, 2021).¹³² Em linhas gerais, o Teorema de Bayes diz respeito a consideração da probabilidade de uma hipótese ser o caso, dada certas condições e evidências disponíveis. Em sua fórmula simples:

$$P(h|e) = \frac{P(e|h) \cdot P(h)}{P(e)}$$

Ou de forma mais estrita:

$$P(h|e) = \frac{P(e|h) \cdot P(h)}{P(e|h) \cdot P(h) + P(e|\sim h) \cdot P(\sim h)}^{133}$$

Entende-se que: P é a probabilidade, h uma hipótese e e a evidência ou os dados disponíveis. Segue-se que, $P(h|e)$ é a probabilidade de h dado e , isto é, a probabilidade de h ser o caso se assumirmos e . Diz-se que e corrobora h se $P(h|e) > P(h)$, pois e torna h mais provável do que seria a probabilidade de que h sem e (GODFREY-SMITH, 2021). Ademais, chama-se $P(h)$ de probabilidade anterior (a probabilidade de que algo seja o caso), $P(h|e)$ de probabilidade posterior (a probabilidade de que algo seja o caso após o exame das evidências) e $P(e|h)$ de verossimilhança (ou, dado uma evidência, o quão provável de que algo é o caso). Assim, $P(\sim h)$ é a negação de h e $P(e|\sim h)$ é se dado a evidência o quão improvável é de que algo é o caso. Aqui o valor em $P(e|\sim h)$ é o oposto de em $P(e|h)$ (GODFREY-SMITH, 2021; PINKER, 2021).

Considere o seguinte exemplo. Suponha que 1% dos indivíduos de uma dada população possuam um tipo de condição congênita severa x . A hipótese de que em uma população certas pessoas possuem essa condição é h . O exame específico para detecção dessa condição que as pessoas realizam, que pode vir a dar positivo, constitui a evidência e . A

¹³² Alguns autores consideram a IME incompatível com o bayesianismo. Para saber sobre esse debate, cf. Henderson, 2014.

¹³³ Pode-se substituir a soma de $P(e)$ pela seguinte equação: $P(e) = \sum_{l \in h} P(e|l) \cdot P(l)$. Esta é uma função da hipótese l , dado a verossimilhança para cada l , dado uma evidência e (PAULINO et al, 2018, p. 10).

primeira informação que pode ser retirada é de que $h = 1\%$ e $\sim h = 99\%$. A primeira hipótese consiste que em uma dada população certos indivíduos possuem um tipo de condição congênita severa x . A segunda hipótese $\sim h$ expressa que em uma dada população certas pessoas *não* possuem a condição congênita severa x . Considere que a chance de o exame dar o resultado positivo, dado que a pessoa realmente tenha a condição congênita severa x , é de 90% e de que a sua negação, isto é, um caso de alarme falso seja de 9.5%. Ou seja, $P(e|h) = 90\%$ e $P(e|\sim h) = 9,5\%$. Por meio do Teorema de Bayes calcula-se que:

$$P(h|e) = \frac{(0.9) \cdot (0.1)}{[(0.9) \cdot (0.1) + (0.95) \cdot (0.99)]}$$

O resultado giraria em torno de 8,7%, deste modo, a probabilidade de o resultado do teste ser positivo seria confirmado pela baixa porcentagem. Em outras palavras, a probabilidade de o indivíduo possuir a condição congênita severa x , dado a realização do exame, é baixa.

Não é o meu intuito discutir de forma exaustiva e estrita as virtudes e dificuldades que a interpretação bayesiana da IME possui. Considero que o uso aqui feito da IME independe do endosso de uma interpretação bayesiana, ou de qualquer outro formalismo probabilístico. No entanto, com isso não quero dizer que se deva excluir esta ou qualquer outra interpretação formal da IME. Creio que a interpretação bayesiana da IME pode ser utilizada de forma frutífera para os casos inferenciais de atribuição de consciência aos outros animais.

3.2.2 Correlatos Neurais, Marcadores e Constituintes da Consciência

Seguindo Matthias Michel e Hakwan Lau (2020), devemos elucidar os conceitos de¹³⁴: correlatos neurais da consciência (NCC, sigla em inglês para *Neural Correlates of Consciousness*)¹³⁵, os marcadores de consciência e os constituintes da consciência, para explorar adequadamente a abordagem epistêmica.

Os NCC's constituem atualmente um dos temas de pesquisa centrais para os cientistas da consciência e, em menor grau, para os filósofos. Podemos definir um NCC como “[...] o conjunto mínimo de eventos e mecanismos neuronais suficientes conjuntamente para

¹³⁴ Novamente, o intuito desta subseção não é exaurir a vasta literatura sobre o assunto, mas distinguir os conceitos de modo a torná-los claros para o prosseguimento do texto. Para um tratamento conceitual rigoroso e extenso sobre o que é um NCC, cf. Chalmers, 2010a.

¹³⁵ Utilizarei a sigla em inglês “NCC” por estar relativamente bem estabelecida na literatura em português.

tornar consciente uma percepção específica.”¹³⁶ (KOCH, 2004, p. 16). Em outras palavras, NCC’s são eventos e mecanismos neuronais que quando tomados em conjunto são suficientes para um estado se tornar consciente. Se uma determinada base neuronal A for identificada como um NCC, quando ativada, ela deve ser capaz de tornar consciente um estado mental que antes era inconsciente (MICHEL e LAU, 2020).

Em um artigo recente, Andreas Nieder e colegas (2020) realizaram um experimento com corvos (*Corvus corone*), um animal conhecido por ter uma cognição sofisticada, em busca de um NCC. Os pesquisadores treinaram dois corvos com o objetivo de “relatarem” a presença ou a ausência de um dado estímulo visual sob um determinado limiar com regras bem estabelecidas. Durante a execução da tarefa, era registrada a atividade neuronal¹³⁷ do nidopálio caudolateral (NCL), região considerada, em certa medida, organizacionalmente análoga ao neocórtex dos mamíferos, embora não possua uma estrutura laminar como o do neocórtex (GÜNTÜRKÜN, 2005; GÜNTÜRKÜN e BUGNYAR, 2016; STACHO et al, 2020). A tarefa era iniciada com um sinal e em seguida era realizado um intervalo curto de 600ms. Posteriormente, era fornecido ou não um estímulo visual aos animais. Os estímulos eram apresentados de forma a serem ou não facilmente perceptíveis. Um período de *delay* era realizado em seguida. Logo após, os pesquisadores forneciam um sinal (*cue*) informando qual o tipo de ação motora os corvos deveriam realizar em seu “relato” acerca da percepção ou não do estímulo que foi ou não apresentado. O sinal (*cue*) funcionava de modo que os corvos não pudessem preparar de antemão a resposta a ser dada. Os pesquisadores classificaram em três categorias diferentes cada forma de apresentação do estímulo visual: a supraliminar, o próximo do limiar e a sem estímulo. O estímulo visual supraliminar era de alta intensidade e o próximo do limiar era de baixíssima intensidade (Figura 12). Foi registrado que durante a apresentação do estímulo, os neurônios do NCL respondiam em sua maior parte de acordo com a intensidade exibida pelo estímulo e menos de acordo com o relato do animal.

No entanto, durante o período posterior ao *delay*, a maior parte dos neurônios do NCL respondiam de acordo com o relato do animal, e não mais com a intensidade do estímulo anteriormente percebido. A atividade neuronal neste período mudava em decorrência daquilo o que o corvo experienciava subjetivamente. Os neurônios respondiam de forma parecida quando o corvo respondia “sim” havendo o estímulo com alta e baixíssima intensidade, e “sim” na ausência do estímulo (alarme falso). O mesmo ocorria com o inverso, a atividade neuronal do

¹³⁶ “[...] *the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept.*”

¹³⁷ Especificamente o registro de célula única de 480 neurônios, sendo 306 do corvo G e 174 do corvo O. Nieder et al, 2020, p. 1627.

“não” quando o estímulo estava próximo do limiar correspondia com o do “não” da percepção da ausência do estímulo. Como os pesquisadores sugeriram, isso provavelmente indicaria a correlação entre a atividade neuronal do NCL com a experiência subjetiva do corvo, e não o mero registro do estímulo fornecido.

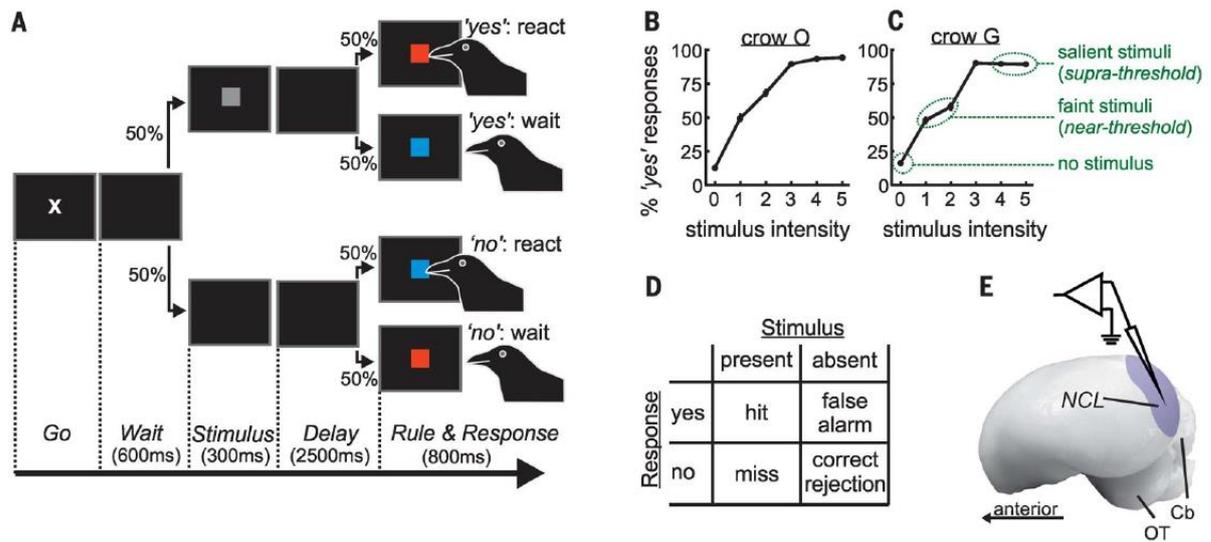


Figura 12 – Esquema geral do experimento de Nieder e colegas. Tradução da esquerda à direita: A: ‘sim’: reagir, ‘sim’: esperar, ‘não’: reagir, ‘não’: esperar. Comece, Aguarde, Estímulo, Delay e Regra e Resposta. B: Corvo O, % de respostas ‘sim’, intensidade do estímulo, estímulo saliente (supraliminar), estímulo fraco (próximo do limiar), sem estímulo. D: Estímulo, presença, ausência, Respostas, sim, acerto, alarme falso, não, erro, rejeição correta. E: Anterior. Nieder et al, 2020, p. 1626. Adaptado.

Os constituintes da consciência, por sua vez, são estados, eventos neuronais, ou algum outro tipo de estado físico qualquer, que são *idênticos* com a própria consciência (MICHEL e LAU, 2020, p. 2). Isto é, para um evento E_n ser considerado um constituinte da consciência, deve não somente ser suficiente, mas necessário para a consciência. Em outras palavras, E_n é um constituinte da consciência se e somente se ele é suficiente e necessário à consciência, ou seja, em todos os mundos possíveis E_n é idêntico à consciência. Da mesma forma que H_2O constitui a água e CO_2 constitui o dióxido de carbono, E_n deve constituir a consciência. Note que a possível existência de um constituinte da consciência E_n faria cair por terra a hipótese da múltipla realizabilidade da consciência, uma vez que a realização da consciência dependeria necessariamente de E_n .

Não sabemos se realmente existe um constituinte da consciência do mesmo modo que sabemos que a água e o dióxido de carbono dependem exclusivamente dos seus compostos químicos. Mesmo que exista, dificilmente conseguiremos conhecê-lo, dadas as próprias limitações da ciência. A concepção da existência de E_n é, portanto, altamente controversa. Por

estes motivos, as pesquisas sobre as bases neuronais da consciência concentram os seus esforços nos NCC's e em menor grau nos marcadores de consciência (MICHEL e LAU, 2020).

Por fim, os marcadores de consciência consistem em evidências¹³⁸, indicadores, capacidades e disposições que, por estarem supostamente relacionados à consciência, sinalizam a presença dela em um dado sistema cognitivo (BAYNE e SHEA, 2020; MICHEL e LAU, 2020). A natureza dos marcadores de consciência pode ser comportamental, neurofisiológico, estrutural ou mesmo tarefas que são facilitadas ou que dependam da consciência para serem executadas. Segundo esta visão, um NCC, como o descrito no exemplo do corvo, poderia ser considerado um marcador. Na medida em que um sistema cognitivo *S* apresente um NCC, isto indicaria que *S* possuiria algum tipo de experiência consciente, por mais limitado que o espectro do NCC fosse (MICHEL e LAU, 2020, p. 3). Claro que o inverso não pode ser dito. Um NCC pode ser um marcador de consciência, mas nem todo marcador de consciência é um NCC. Como veremos a seguir, existem outros tipos de marcadores que indicam confiavelmente a presença da consciência. Note também que um marcador de consciência pode não ser suficiente, menos ainda necessário, para que a consciência seja realizada. De forma similar, os sinais de ser líquido, transparente e inodoro podem indicar a provável presença de água, mas não são suficientes e nem necessárias para a realização de H₂O. Assim, como um conjunto de sintomas fornecem evidências da presença da doença, mas não são suficientes e nem necessários para a existência dela.

3.3 A Estratégia dos Marcadores de Consciência

Como forma de responder os problemas da distribuição e da mensuração, as propostas epistêmicas buscam se esquivar do comprometimento total com teorias prévias da consciência humana, como fazem os proponentes da abordagem teórica. Além disso, os teóricos da abordagem epistêmica também se afastam de possíveis mecanismos inatos que supostamente seriam capazes detectar os estados conscientes dos animais. Para que tal empreendimento obtenha sucesso, a estratégia que a maior parte das teses epistêmicas utilizam em comum é o que chamo de “*estratégia dos marcadores de consciência*” (EMC). De forma estrita, a EMC expressa que:

¹³⁸ Entendo aqui evidência de forma ampla, como algo que fornece justificativas racionais à hipótese *H* acerca de *P*. Para fins desta dissertação, não será necessário o comprometimento com nenhuma concepção mais estrita do conceito de “evidência”. Mantereí, portanto, uma postura neutra no que concerne o debate entre as três principais posições atuais: o bayesianismo, o frequentismo e o *likelihoodism*. Para maiores informações, cf. Godfrey-Smith, 2021, cap. 12; Sober, 2009, cap. 1.

Dado o funcionamento das leis da natureza em nosso mundo atual, os marcadores de consciência, por estarem de alguma forma associados à consciência, constituem evidências que justificam a melhor hipótese acerca da atribuição de consciência a sistemas cognitivos.

A abordagem epistêmica, deste modo, assegura a sua posição por meio dos marcadores de consciência. São eles que servirão como evidências capazes de *justificar* a hipótese *x* ou *y* acerca da presença dos estados mentais conscientes em um determinado animal.¹³⁹ Note que a EMC não diz que os marcadores de consciência seriam necessários ou mesmo suficientes para que haja consciência no animal em questão. Eles seriam apenas indicadores que, por estarem associados à consciência, justificam a IME em relação à hipótese *H* sobre a consciência no dado animal. Assim, caso um determinado animal *S* não apresente um marcador de consciência M_1 , disto não se segue que *S* não possua consciência. A capacidade de relato verbal, por exemplo, pode contar como um forte indicador para atribuição de consciência. No entanto, a carência do relato verbal claramente não implica na ausência da consciência no animal. Ou seja, pode ser o caso que *S* não apresente M_1 , mas apresente outros marcadores (M_2 , M_3 , $M_n...$) que indiquem nele a presença da consciência. Assim, M_1 não constituiria uma condição necessária para a presença da consciência. Também pode ser o caso de que *S* apresente M_1 , mas que por algum motivo não seja um animal dotado de consciência, isto é, ser um alarme falso. Neste caso M_1 não constituiria uma condição suficiente para a presença da consciência. As I. A. dos smartphones que são projetadas para conversarem verbalmente com o seu usuário claramente não são conscientes, mesmo que possua o marcador de relato. A pergunta que deve ser levantada é a seguinte: como podemos então assegurar que um marcador M_n confiavelmente justifique a IME de que *S* possui consciência? Em outras palavras, de que modo uma abordagem epistêmica pode ser capaz de indicar de forma segura a presença da consciência em um determinado animal? Diante de tais questionamentos, cada proposta da abordagem epistêmica fornecerá uma resposta própria. Apresentarei algumas delas nas próximas seções.

Mesmo os autores que formularam teorias robustas da consciência humana, como Baars (1988, 2005a) e Tye (1995), adotaram alguma versão da abordagem epistêmica em relação à consciência animal (BAARS, 2005b; TYE, 2017, 2018). Apesar da maior parte das teses epistêmicas possuírem em comum a EMC, cada uma a utiliza de forma distinta. Por meio da EMC, alguns autores consideram, por exemplo, a consciência como um tipo natural (*natural*

¹³⁹ Ressalto, novamente, que aqui me restrinjo a somente os casos dos animais. No entanto, como Shevlin (2021) também aponta, esta estratégia pode ser estendida a casos como os das I. A., ou mesmo o de humanos em estado vegetativo, como Shea e Bayne (2010) defendem.

kind), cuja presença pode ser identificada pelo conjunto (*cluster*) de marcadores supostamente associados à consciência (BAYNE e SHEA, 2020; SHEA e BAYNE, 2010; SHEA, 2012). Outros buscam marcadores de consciência neurofisiológicos e anatômicos em humanos, como algumas estruturas do córtex e do tálamo, e comportamentais com alto poder de discriminação que associem estados perceptuais à memória, para encontrar estruturas homólogas em outros animais (EDELMAN et al, 2005).¹⁴⁰ A EMC pode até mesmo ser utilizada de forma conjunta com a abordagem teórica. Como, por exemplo, utilizando os marcadores de consciência como meio de especificar quais os sistemas consumidores, ou requisitos cognitivos, uma teoria da consciência humana deve possuir para indicar se um determinado animal possui ou não consciência (SHEVLIN, 2021).

A origem da abordagem epistêmica remonta à etologia cognitiva. Durante grande parte do século XX, os estudos acerca da mente dos animais, e até mesmo dos humanos, estava eclipsada pelos estudos da análise comportamental. Falar sobre a mente, a consciência ou qualquer outro estado interno era considerado um tabu. Ou os estados mentais, bem como os conscientes, eram totalmente eliminados do vocabulário científico, ou então eram desconsiderados em prol dos estudos comportamentais. Com a revolução cognitiva da segunda metade do século XX, as pesquisas acerca da cognição humana voltaram à tona (BERMÚDEZ, 2020). Neste contexto, Donald R. Griffin (1992, 2002) foi um dos pioneiros da retomada dos estudos da cognição, da intencionalidade e, sobretudo, da consciência entre os animais, criando o programa de pesquisa denominado de “etologia cognitiva” (ALLEN e BEKOFF, 1997; ANDREWS, 2020a, p. 86). Em sua metodologia, Griffin já sinalizava que a utilização de um corpo de evidências robustas poderia servir como base para sabermos se um animal é consciente. Em suas palavras:

Embora nenhuma única evidência forneça uma prova absoluta da consciência, a acumulação destas evidências fortemente sugestivas aumenta significativamente a probabilidade de que alguns animais experenciam ao menos formas simples de pensamentos e emoções conscientes.¹⁴¹ (GRIFFIN e SPECK, 2004, p. 5)

¹⁴⁰ Para uma crítica à posição da homologia, cf. Key et al, 2021.

¹⁴¹ “Although no single piece of evidence provides absolute proof of consciousness, this accumulation of strongly suggestive evidence increases significantly the likelihood that some animals experience at least simple conscious thoughts and feelings.”

Além de Griffin e Baars, outros cientistas sustentaram alguma versão própria da abordagem epistêmica (cf. EDELMAN e SETH, 2009; EDELMAN et al, 2005; SETH et al, 2005). Ademais, outros filósofos além de Tye adotaram o modelo epistêmico, como Andrews e a abordagem dos marcadores dinâmicos (2020a, pp. 86-87), Bayne e Shea com a estratégia do tipo natural (2020; SHEA, 2012),¹⁴² Paul Thagard com o procedimento de atribuição (2021) e Shevlin com a abordagem teórica modesta (2021).

Isto posto, analisarei a seguir duas posições epistêmicas importantes, a saber, a chamada “abordagem padrão”, popular entre o senso comum (SHEA e BAYNE, 2010) e a abordagem newtoniana de Tye (2017, 2018). Em seguida, defenderei a abordagem teórica leve, posição que considero ser atualmente a mais adequada (BIRCH, 2020a). Por fim, tentarei responder a possíveis objeções.

3.4 A Abordagem Padrão

A chamada “abordagem padrão” foi originalmente apresentada e atacada por Shea e Bayne (2010) em um artigo sobre atribuição de consciência nos casos de humanos em estado vegetativo. No entanto, creio que essa abordagem também possa ser transportada aos casos dos animais. Por ela soar à primeira vista intuitiva, acredito ser importante refletirmos a respeito para, com os devidos motivos, abandoná-la. Os autores a nomearam desta forma, pois eles a consideram dominante no debate acerca dos humanos em estado vegetativo. Apesar de não a considerar dominante no caso dos animais não mudarei a sua nomenclatura. Acredito que a manutenção da nomenclatura evitará confusões. Ademais, mantereí a nomenclatura pois acredito que aos olhos do senso comum essa abordagem ainda permanece razoável. Claro, isso é uma percepção minha que para ganhar robustez deveria ser empiricamente testada.

De acordo com Shea e Bayne (2010), a abordagem padrão (AP) sustenta que:

[...] a atribuição de consciência em casos contestados, como os de humanos em estado vegetativo, devem ser somente governados por esses marcadores pré-teóricos que utilizamos para atribuímos consciência uns aos outros, como, a reportabilidade e as várias formas de comportamento voluntário.¹⁴³
(SHEA e BAYNE, 2010, p. 463)

¹⁴² Bayne e Shea (2020) também consideram a sua própria abordagem um modelo da abordagem teórica leve.

¹⁴³ [...] *the ascription of consciousness in contested cases such as vegetative state should be governed solely by those pre-theoretical markers that we use to ascribe consciousness to each other, namely, reportability and various forms of voluntary behaviour.*

Podemos ampliar os casos contestados, isto é, os casos em que ainda existe um importante debate a respeito da presença da consciência, a, pelo menos, os humanos em estado comatoso, bebês recém-nascidos, I. A. e, claro, os animais. Aquilo o que os autores chamam na citação acima de “marcadores pré-teóricos”, nada mais são que tipos de marcadores de consciência de natureza comportamental, cujos proponentes da AP consideram estar fortemente associados à consciência (SHEA e BAYNE, 2010, p. 460). Para a AP, são esses marcadores que funcionarão como o “*padrão ouro*” para a atribuição de consciência. Seguindo esse raciocínio, se o animal apresentar um dos marcadores comportamentais do padrão ouro, isto justificaria a atribuição de consciência a ele. Segundo Shea e Bayne (2010), dois argumentos são utilizados para a sustentação dessa abordagem. O primeiro é o argumento da reportabilidade e o segundo da volição. Por razões de economia, abordarei somente o primeiro argumento que é sem dúvidas o mais problemático.¹⁴⁴

Segundo a AP, um dos marcadores mais confiáveis que deve ser utilizado como padrão ouro é a capacidade do sistema produzir um relato (*report*) (LEDOUX, 2019; NACCACHE, 2018).¹⁴⁵ Existem ao menos duas formas de relatos: o “introspectivo” (*introspective report*) e o “ambiental” (*environmental report*). O primeiro é um relato sobre os próprios estados experienciais que ocorrem em sua mente. Por exemplo, o pensamento acerca da experiência da dor que sinto no momento. O segundo, por sua vez, é o relato sobre o seu próprio ambiente, incluindo o seu entorno e os seus estados corporais (SHEA e BAYNE, 2010, p. 464). Por exemplo, o relato sobre a presença de um computador a minha frente. Segundo esta versão do argumento, é somente via o relato introspectivo e/ou ambiental que poderíamos atribuir confiavelmente consciência a um animal, pois seria o relato que forneceria acesso aos estados mentais conscientes dele.

Seria, portanto, a capacidade de relatar as experiências do que ocorre internamente ou em seu ambiente que contaria como o marcador de consciência M_R . Em outras palavras, é por meio do relato comportamental R que possuímos acesso aos estados mentais conscientes de S . Se S não é capaz de fornecer R , então não temos acesso aos estados conscientes de S (assumindo que S possua consciência). Logo, não podemos atribuir de forma confiável estados conscientes a S . Por exemplo, se o animal não for capaz de relatar, por meio do comportamento, os seus estados experienciais conscientes, como, por exemplo, a dor em uma de suas patas, ou o que está acontecendo em seu entorno, como a percepção visual de uma árvore balançando a

¹⁴⁴ Cf. o artigo original para um tratamento sobre o argumento da volição, Shea e Bayne, 2010.

¹⁴⁵ “*The fact that animals can only respond nonverbally means there is no other response to help distinguish conscious from nonconscious processes.*” (LEDOUX, 2019, p. 320)

sua frente, não conseguirem obter acesso a tais estados. Sendo assim, não poderíamos justificadamente atribuir consciência ao animal em questão.

Como neste caso o padrão ouro é o marcador comportamental de relato M_R , na ausência dele não seria possível indicar a presença da consciência no animal em questão. Existem ao menos três problemas com este argumento. O primeiro é que ele parece exigir que o animal seja autoconsciente. O segundo requer injustificadamente a capacidade de relatar. E terceiro, o argumento se vale de apenas um único marcador que julgará a presença da consciência.

Na medida em que se requisita a capacidade de relato dos seus próprios estados conscientes, a abordagem padrão acaba se comprometendo com alguma forma de intelectualismo. Se para julgarmos se um animal é ou não consciente for necessário a apresentação de M_R , estamos no fundo demandando que o animal possua a capacidade de ter autoconsciência acerca de seus estados mentais de primeira ordem. Por exemplo, se o animal não for capaz de pensar, ou ter o conceito necessário, sobre o seu estado mental de dor, isto é, se ele não for consciente da sua experiência de dor, ele não conseguirá relatar aquilo o que está ocorrendo consigo mesmo. Ou seja, para que o animal S apresente M_R , S deve ser consciente de seu estado mental de nível inferior E , para que consiga, assim, fornecer um relato comportamental sobre E . Ora, tal exigência acaba por comprometer a AP com alguma forma de intelectualismo. Claro, com isso não estou querendo dizer que quando um animal exhibe um comportamento C , que é usualmente associado com a dor, ele estaria automaticamente autoconsciente do seu estado mental de dor. Expressar que toda forma de C implica em autoconsciência do estado mental referente E , seria se comprometer com alguma tese de ordem superior da relação entre C e E . Não é este o meu intuito. Exibir um comportamento que expressa dor não seria a mesma coisa que exibir um *comportamento de relato* da dor. O primeiro tipo de comportamento não implica na autoconsciência da dor, já o segundo sim. Não vejo motivos para se exigir algo tão complexo e controverso quanto à autoconsciência.

No entanto, penso não haver ganhos teóricos mesmo entre aqueles que porventura endossem um tipo de intelectualismo. O motivo é que eles se comprometeriam tacitamente com uma teoria com uma concepção prévia acerca da consciência. Isto vai de encontro a uma das principais assunções da abordagem epistêmica, a saber, a possibilidade de não se comprometer previamente com uma teoria da consciência. Além disso, eles estariam endossando uma teoria que é controversa na literatura (BLOCK, 2011b, 2019; CARRUTHERS, 2019; KOZUCH, 2014). Um dos problemas que esse tipo de teoria recai seria na determinação do conteúdo de

uma experiência consciente. Quem neste caso determinaria o conteúdo de um certo estado consciente? O estado mental de primeira ordem ou o de ordem superior? Considere o seguinte exemplo retirado e levemente alterado de Birch (2021). Suponha que estou neste momento percebendo diante de mim uma mesa vermelha. No entanto, por uma razão qualquer, penso que estou percebendo-a como sendo azul. Diante deste fato, qual seria a cor que estou conscientemente experienciando no momento, vermelho ou azul? Digamos que eu afirme “vermelho”. Neste caso, o conteúdo de ordem superior se torna dispensável ao conteúdo da própria experiência, uma vez que ela não seria preciso ao conteúdo da consciência. Por outro lado, caso eu diga “azul”, o conteúdo de ordem superior se torna relevante, porém ele entraria em conflito com o conteúdo da experiência (vermelho) que instancia o conteúdo dos nossos conceitos (BIRCH, 2021, pp. 6-7).

Mesmo para os humanos M_R se revela demasiado exigente. Pense, por exemplo, nos casos de bebês com poucos meses de vida e de recém-nascidos. Parece ser contraintuitivo que recém-nascidos possuam inatamente os conceitos necessários para se auto referir de forma consciente. Porém, ao que tudo indica, eles são seres conscientes (FERREIRA, 2017; LAGERCRANTZ e CHANGEUX, 2009).¹⁴⁶ O mesmo poderia ser dito para os casos intuitivos dos animais, como os cachorros, aos quais não hesitamos em atribuir consciência, mas relutamos em atribuir autoconsciência.

Ora, com isso não estou dizendo que estados mentais de ordem superior e formas gerais de metacognição¹⁴⁷ sejam dispensáveis aos estudos sobre a consciência animal, nem tampouco que M_R não seja um marcador de consciência relevante. O que exponho é que a capacidade de metacognição pode servir como um importante indicador para saber se um animal possui consciência,¹⁴⁸ mas ele não deve ser utilizado como o padrão ouro para tal.¹⁴⁹ O risco de se obter falsos negativos utilizando M_R como a medida central poderia ser alto. Isto incluiria os casos que parecem ser bem intuitivos, como os de bebês, recém-nascidos e animais filogeneticamente próximos de nós, como os chimpanzés. Como veremos nas próximas seções, existem outros marcadores de consciência que são tão, ou mais, importantes que M_R .

¹⁴⁶ Com isso não estou dizendo que bebês ou recém-nascidos não possuam certos conceitos inatos. Pelo contrário, me parece razoável que alguns conceitos o sejam, como o de “objeto” e “número”. Evidentemente, os bebês *não são* uma tábula rasa (BERENT, 2020; DEHAENE, 2021).

¹⁴⁷ Para uma introdução acessível às pesquisas sobre metacognição em geral, cf. Fleming, 2021. Em relação aos questionamentos acerca da metacognição em animais, cf. Smith et al, 2014; Proust, 2018.

¹⁴⁸ Shea e Heyes (2010) sugerem que a metamemória, por exemplo, poderia servir como uma boa evidência da presença da consciência em animais.

¹⁴⁹ Tye (2017, pp. 20-21) também compartilha da mesma preocupação.

Consideremos agora o segundo problema que diz respeito ao requerimento injustificado da capacidade de relatar. À primeira vista pode parecer ser correto que se exija o relato comportamental como meio de ter acesso aos estados conscientes do animal. No entanto, essa me parece ser uma exigência equivocada. Como vimos, a AP apresenta dois tipos de relatos: o introspectivo e o ambiental. O primeiro diz respeito à capacidade do animal fornecer um relato acerca dos seus próprios estados mentais. Porém, me parece que para que o animal seja capaz de fornecer o relato introspectivo, ele deve possuir autoconsciência, na medida em que ele deve ter ciência do seu estado mental a ser relatado. Entretanto, isto resultaria no problema apresentado anteriormente. Os bebês com poucos meses de vida, recém-nascidos e muito dos animais seriam automaticamente considerados incapazes de fornecer o relato. Eles seriam considerados seres destituídos de consciência (SHEA e BAYNE, 2010, p. 465).

Em contrapartida, o relato ambiental suscita outros problemas. A incapacidade de fornecer um relato acerca do seu entorno não deve implicar na ausência de consciência do animal. Como vimos no capítulo 1, as pessoas que sofrem da síndrome do encarceramento são quase, ou totalmente, incapazes de relatarem aquilo que se passa a sua volta (ou mesmo o que se passa internamente) através do comportamento. Se porventura formos nos guiar apenas pelo relato, diversos casos de pessoas conscientes que se encontram incapazes de relatar seriam consideradas inconscientes. Por exemplo, utilizando a ressonância magnética funcional (fMRI),¹⁵⁰ Owen e colegas (2006) alegaram terem se comunicado com um paciente em estado vegetativo recorrendo apenas ao uso de imagens mentais. Dentro do contexto da AP, este paciente seria considerado inconsciente, pois a sua “comunicação” não foi feita por meio da exibição do comportamento. Da mesma forma poderíamos transportar essas reservas aos casos dos bebês com poucos meses de vida, dos recém-nascidos e dos animais. Não é porque o animal não pode fornecer um relato em forma de comportamento que ele seria um ser inconsciente. Assim como um indivíduo com síndrome de encarceramento que está impossibilitado de exercer qualquer tipo de comportamento pode estar consciente de tudo o que se passa internamente e em seu entorno, um animal também pode se encontrar em uma condição análoga ao do paciente. Ou seja, consciente, mas incapaz de fornecer um relato.

Por fim, a AP se equivoca ao utilizar somente um único marcador em seu julgamento sobre se um animal *S* é ou não consciente. Tendo isso em vista, um portador da síndrome do encarceramento seria considerado inconsciente, o que consistiria em um caso de

¹⁵⁰ Sobre o funcionamento dos mecanismos e procedimentos utilizados em imageamento cerebral cf. Backer, 2022; Passingham e Rowe, 2016.

falso negativo. Ademais, poderíamos conceber o seguinte cenário. Suponha que João é o cientista mais genial e habilidoso de seu tempo. Como o projeto último de sua vida, ele fabrica um robô que visto externamente é incrivelmente parecido com a aparência de um ser humano comum. Todos os seus *outputs* também são idênticos ou muito semelhantes aos dos humanos. Quando perguntado sobre o que lhe ocorre internamente ou em seu entorno, o robô seria capaz de fornecer um relato indiscernível daquele fornecido por uma pessoa comum. Por exemplo, o robô poderia relatar a dor de cabeça que sente neste momento e a percepção do vento movendo as cortinas da janela de seu quarto.

Poderíamos também retomar o teste de Turing. Imagine um computador tal como a máquina de Turing, mas com a qual não tivéssemos contato visual (a máquina estaria em um quarto adjacente ao nosso). Essa máquina seria capaz de produzir relatos verbais, se não idênticos ao menos muito semelhantes aos dos humanos sobre seus estados internos e sobre o que acontece no ambiente. Quando perguntada sobre o que lhe ocorre internamente ou em seu entorno, a máquina seria capaz de fornecer um relato indiscernível daquele fornecido por um humano comum. Por exemplo, ela poderia relatar uma dor na perna, raiva, ansiedade etc. Além disso, ela também poderia relatar a percepção auditiva de um zunido feito por uma mosca. Em ambos os exemplos supracitados, o robô e a máquina satisfariam M_R , mas claramente não seriam entidades genuinamente conscientes. Ou seja, ambos seriam casos de alarme falso. A AP estaria, portanto, sujeita a erros triviais, atribuindo consciência a sistemas que não possuem e excluindo daqueles que a tem.

Isto posto, considero plausível que se abandone a AP no que se refere à atribuição de consciência a animais. A AP, de acordo com os argumentos apresentados acima, não consistiria em uma boa resposta ao problema da mensuração. Ela se vê às voltas com uma série de problemas que a tornam uma metodologia imprecisa. Mesmo em seu alvo original, a saber, os casos humanos controversos, ela falharia em sua proposta (SHEA e BAYNE, 2010).

3.5 A Abordagem Newtoniana de Tye

Outra proposta epistêmica relevante é a que Birch (2020a) denominou de a “abordagem teórica neutra”. Segundo o autor, esta abordagem seria bastante popular na literatura sobre dor em animais (BIRCH, 2020a). A ideia aqui seria manter uma atitude completamente afastada e neutra em relação a qualquer tipo de comprometimento teórico e metafísico sobre a consciência. Nesta seção, tratarei de uma versão particular da abordagem teórica neutra, a que chamo de *abordagem newtoniana*. Essa abordagem pode ser exemplificada

por meio da posição de Tye.¹⁵¹ Embora possua uma teoria própria da consciência (TYE, 1995), Tye propõe uma abordagem teórica neutra de orientação newtoniana para a atribuição de consciência a animais (TYE, 2017, 2018).¹⁵² Ele o faz através da introdução da chamada “Regra de Newton”.

Para uma melhor compreensão da Regra de Newton, devemos antes entender aquilo que Tye (2017, pp. 12-17) denomina de “Visão Simples”. Segundo Tye, a Visão Simples afirma que “uma criatura está consciente no tempo t se e somente se ele está passando por uma ou mais experiências em t .”¹⁵³ (TYE, 2017, p. 12). Ademais, “*um estado mental é consciente se e somente se ele é uma experiência.*”¹⁵⁴ (TYE, 2017, p 12, grifo meu). Para Tye, uma experiência é um estado mental M de maneira que há algo como ser x para um sujeito S . Portanto, M é um estado consciente se e somente se existe algo como ser x para S que realiza o estado M (TYE, 2017, pp. 12-13). Como Carruthers (2019, p. 54) bem aponta, Tye identifica a consciência com a experiência. Ou seja, a consciência é nada mais que a própria experiência. Assumindo que percepção e experiênci são a mesma coisa, isto implicaria na tese segundo a qual todas as formas de percepção são estados conscientes (CARRUTHERS, 2019, p. 54).

Como forma de tornar claro, considere o célebre exemplo levemente modificado do motorista de Armstrong (1993). Imagine as experiências visuais de um motorista distraído dirigindo em uma longa e monótona estrada (TYE, 2017, p. 14). Durante o percurso, o motorista está concentrado na realização de outras tarefas que não a da própria direção, como falar no celular sobre um assunto de trabalho, a mastigação de uma goma de mascar e a introdução de uma garrafa no porta copo do carro. Diante desta situação, diríamos que apesar do motorista possuir uma experiência visual, ela seria *inconsciente*. No entanto, como Tye (2017, pp. 14-15) salienta, afirmar a existência de uma experiência inconsciente tornaria inconsistente a Visão Simples, pois seria admitido uma experiência de uma experiência que não é uma experiência. Colocando de outra forma, neste cenário o motorista teria uma experiência inconsciente, porém como a Visão Simples identifica a experiência com a consciência, seria inconsistente o

¹⁵¹ Embora difiram da hipótese de Tye, considero a posição de Thagard (2021) e Fischer (2021) pertencentes à abordagem newtoniana. Hume (2009) me parece ter antecipado essa posição formulando um argumento semelhante. Cf. o 1º cap. desta dissertação.

¹⁵² Em seu livro de 1995, Tye formulou uma teoria representacionista de primeira ordem da consciência, a chamada “PANIC”. Ela consiste em um modelo metafísico da GWT. Todavia, em seu livro mais recente de 2021, Tye realizou uma modificação curiosa em sua tese. Ele combinou o pampsiquismo com o representacionismo, criando o “representacionismo pampsiquista” (TYE, 2021, 2022).

¹⁵³ “[...] *a creature is conscious at time t if and only if it is undergoing one or more experiences at t .*”

¹⁵⁴ “[...] *a mental state is conscious if and only if it is an experience.*”

motorista ter a experiência que é uma experiência que não é uma experiência (TYE, 2017, p. 14). Ou formalizando, sendo E = experiência, T = ter e m = motorista:

$$\exists x (Ex \wedge Tmx \wedge \neg Ex)$$

Lê-se a fórmula acima como: há um x tal que, x é uma experiência e o motorista possui essa experiência e essa experiência *não* é uma experiência. Ora, o que se tem aqui é uma contradição, uma vez que o motorista possuiria uma experiência que é ao mesmo tempo uma experiência e a sua negação. Para que a Visão Simples obtenha sucesso, Tye deve se esquivar da contradição que se segue dessa primeira caracterização de sua tese quanto a estados perceptuais inconscientes.

Como resposta, Tye realiza uma alteração na Visão Simples. Para Tye (2017), devemos distinguir um estado mental intrinsecamente consciente (C_1) da atitude do sujeito em relação a esse estado, chamo de C_2 . Nesta concepção, C_2 seria aquilo o que denominei no 1º capítulo de “consciência de criatura transitiva”, isto é, a consciência *de (conscious of)* alguma coisa.¹⁵⁵ Deste modo, o sentido de consciência para a Visão Simples consistiria em C_1 e não C_2 . Ora, segundo este modelo, o motorista seria inconsciente no sentido de C_2 e não em C_1 . O motorista possuiria uma experiência perceptual consciente de primeira ordem, mas não teria consciência de segunda ordem dela (TYE, 2017). Apesar do motorista ter uma experiência perceptual da estrada, ou seja, C_1 , ele não teria ciência (*awareness*) em relação à C_1 . Faltaria ao motorista a consciência em seu sentido C_2 . Assim sendo, os estados perceptuais inconscientes seriam apenas estados que não possuem uma ciência (*awareness*) em relação a eles próprios. Ou seja, a percepção inconsciente, segundo Tye, seriam estados que *não* são acompanhadas de um estado de ordem superior. A Visão Simples, portanto, considera todas as formas de estados perceptuais como sendo conscientes em seu sentido fenomenal (CARRUTHERS, 2019, p. 54).

Isto posto, Tye (2017, 2018) propõe que a melhor forma de atribuir consciência a outros animais é via a aplicação da Regra de Newton. É ela que justificará a IME a respeito da hipótese da presença da consciência no animal em questão. Do mesmo modo que o célebre físico Isaac Newton teria formulado na “Regra 2” que afirma que “*aos mesmos efeitos naturais temos de atribuir as mesmas causas, tanto quanto possível.*” (NEWTON, 2012, p. 185), Tye

¹⁵⁵ Neste cenário, estaríamos identificando um caso particular de consciência de criatura transitiva com um estado consciente de ordem superior. No entanto, existem casos de consciência de criatura transitiva de primeira ordem, como a consciência do livro que se encontra diante de mim. Com efeito, é plenamente debatível se a consciência de criatura transitiva depende ou não de estados mentais conscientes, seja em seu sentido fenomenal ou de acesso.

busca aplicá-lo aos estados conscientes dos outros animais. Em suma, a Regra de Newton pode ser abreviada da seguinte maneira; “mesmo efeito, mesma causa”. A abordagem newtoniana é expressa da seguinte forma por Tye:

Suponha [...] que humanos e animais não humanos engajem no mesmo ou muito próximo comportamento *B* [...], dado o mesmo *noxious stimulus S*. Por que os humanos produzem um comportamento *B*, dado *S*? Porque *S* provoca o sentimento de dor neles, e o sentimento de dor causa *B*.¹⁵⁶ (TYE, 2017, p. 75)

Ele continua:

Indo agora ao caso dos animais não humanos, estou apto a inferir que o sentimento de dor também causa o comportamento *B* neles, a menos que eu tenha um *invalidador (defeater)* àquela inferência, isto é, uma razão para acreditar que a história causal é operativa nestes animais que não podem ser reconciliados com a história causal que anuncie à dor.¹⁵⁷ (TYE, 2017, p. 75)

A Regra de Newton pode ser estruturada da seguinte forma:

- E) Diante do estímulo *G* os humanos exibem um comportamento *C* cuja causa é o estado consciente *E*.
- F) Diante do estímulo *G'* o animal não humano *S* exibe um comportamento *C'*.
Se não existem evidências contrárias que possam invalidar a hipótese da realização de *E'* em *S*, por meio da IME infere-se que:

G) Diante de *G'*, *S* exibiu *C'* cuja causa foi *E'*.

H) Portanto, *S* possui *E'*.

¹⁵⁶ “Suppose as above that humans and nonhumans animals engage in the same or very similar behavior **B** [...], given the same *noxious stimulus S*. Why do humans produce behavior **B**, given *S*? Because *S* elicits the feeling of pain in them, and the feeling of pain causes **B**.”

¹⁵⁷ “Turning now to the case of nonhuman animals, I am entitled to infer that the feeling of pain causes behavior **B** in them too unless I have a **defeater** to that inference, that is, a reason to believe that a causal story is operative in those animals that cannot be reconciled with the causal story that adverts to pain.”

Suponha, por exemplo, que a causa do comportamento flexível C_1 produzido pelo ser humano esteja associada à consciência. Se observarmos C_1 em outros animais, dado a IME, estaríamos justificados a explicar a sua causa através do emprego da consciência. Neste caso, seria mais razoável explicarmos a história causal de C por meio da consciência do que a negarmos como a sua causa, ou postularmos uma causa distinta. Seria, portanto, mais parcimonioso atribuímos a mesma causa em razão do mesmo efeito, do que invocar algo como um estado mental inconsciente, ou negarmos de maneira *ad hoc* que o animal em questão possua estados mentais conscientes. Como Carruthers (2019, p. 53) nota, a ideia básica da Regra de Newton é a navalha de Ockham que, em linhas gerais, pode ser expressa da seguinte forma: não devemos multiplicar entidades ou fatores explanatórios onde não há necessidade.

Com efeito, a mera observação de um animal exibindo um comportamento C , cuja causa nos humanos é associada à consciência, não justifica a inferência de que o animal possui consciência. Isto resultaria naquilo que Birch (2020a) denominou de “credulidade”. Considere, por exemplo, C como um comportamento que em humanos é causado pela consciência e S um animal qualquer. Se toda vez que S apresentasse C inferíssemos que S possui consciência, poderia ser o caso de que a causa de C em S não fosse a consciência, mas algum outro estado qualquer (mental ou não). Isto implicaria na credulidade quanto a essa inferência que atribui consciência a qualquer sistema cognitivo que exibisse C . Desta forma, até mesmo uma máquina programada para somente reproduzir C seria considerada dotada de consciência. A utilização acrítica dessa inferência resultaria na atribuição inadvertida da consciência a animais e outros sistemas cognitivos que, por sua vez, não a possuem. Obviamente não é isso o que Tye busca. Assim, a condição do invalidador (*defeater*) se torna essencial para a formulação da Regra de Newton.

Por exemplo, em humanos os comportamentos de evasão aprendidos (*learned avoidance behaviors*) geralmente possuem a consciência entre as suas causas (BIRCH, 2020a). Grosso modo, um comportamento de evasão aprendido é o que se passa quando um organismo aprende a associar um estímulo nocivo (eventos reais ou potencialmente danosos a um tecido qualquer do organismo) a partir de uma experiência desagradável (BIRCH, 2020a).¹⁵⁸ No entanto, esse tipo de comportamento já foi verificado em ratos que tiveram a sua medula espinhal desconectada do cérebro e em nematoides, como as *Caenorhabditis elegans*, que possuem por volta de 300 neurônios em seus cérebros (ALLEN et al, 2009; BIRCH, 2020a; GRAU, 2017; HA et al, 2010). A princípio, esses seriam casos considerados extremamente

¹⁵⁸ Para maiores informações, cf. Krypotos et al, 2015.

contraintuitivos no que se refere à atribuição de consciência. Como Birch (2020a) nota, isso não quer dizer que os ratos com a medula espinhal desconectada do cérebro ou os nematoides não possuam experiências conscientes de qualquer tipo.¹⁵⁹ Esses exemplos servem apenas para ilustrar a importância do invalidador para a Regra de Newton, pois sem ele trivializaríamos a atribuição de consciência. É o invalidador que impede que a Regra de Newton caia em credulidade, atribuindo inadvertidamente consciência a todos os animais que exibissem *C*, sem que fosse levado em conta possíveis impeditivos, como lesões severas, simplicidade do sistema nervoso central, entre outros.

Por fim, Tye (2017, p. 162) sugere que a flexibilidade comportamental poderia servir como o critério demarcatório que separa os organismos dotados de estados mentais, dentre eles os conscientes, dos que são destituídos de estados mentais. Tye denomina de “teste dos zumbis” tal procedimento. Em linhas gerais, um comportamento flexível é aquele onde, em um dado ambiente, o organismo consegue ajustar a sua resposta face a um estímulo similar ou igual ao anteriormente apresentado (OKASHA, 2020, p. 24). Isto é, responder de diferentes formas a um mesmo estímulo dado um certo contexto (CRUMP et al, 2022, p. 7). Via de regra, o repertório disponível de comportamentos flexíveis de um organismo é resultado de um certo nível de complexificação do sistema nervoso e do processamento cognitivo, como o aprendizado, a capacidade de realizar inferências, a memória entre outros (OKASHA, 2020, p. 24).

Considere a seguinte ilustração. As formigas das espécies *Pogonomyrmex badius* e *Solenopsis saevissima* exibem um comportamento do tipo *necrofórico* (*necrophoric behavior*). Ao encontrarem o corpo de uma irmã operária morta próxima de seu ninho, elas a recolhem e a levam a uma espécie de “pilha de descarte”. Foi descoberto que os corpos em decomposição das formigas mortas produzem uma certa substância química, um ácido graxo, que provoca este tipo de comportamento necrofórico entre as formigas irmãs a sua volta. Dito isto, os pesquisadores buscaram entender se diante de outros cenários essa substância química provocaria o mesmo comportamento necrofórico entre as formigas. Para isso, eles recolheram o ácido graxo dos corpos das formigas mortas e o aplicaram em outros veículos, como papéis e sementes. As formigas que entraram em contato com os papéis e sementes com tal substância química aplicada exibiram o mesmo tipo de comportamento necrofórico, levando-as também à pilha de descartes. No entanto, o resultado obtido mais interessante foi quando os pesquisadores

¹⁵⁹ Embora defendam que os artrópodes possuam consciência, Barron e Klein (2016) negam a presença dela entre os nematoides.

aplicaram a substância nos corpos de formigas vivas. Foi observado que, diante dessas formigas vivas com o ácido graxo aplicado em seus corpos, mesmo elas se locomovendo e apresentando comportamentos típicos, as outras formigas ainda assim exibiam comportamentos necrofóricos. Elas recolhiam as suas irmãs operárias vivas, mesmo que se mexendo, e as levavam à pilha de descarte (WILSON et al, 1958).

Como pode-se notar, essas formigas apresentavam comportamentos inflexíveis, pois mesmo diante de estímulos conflitantes, como uma formiga viva se comportando tipicamente, mas “exalando” o ácido graxo, elas não ajustavam o seu comportamento. Em outras palavras, as outras formigas não ajustavam os seus comportamentos de acordo com o estímulo comportamental observável da formiga viva com o ácido graxo aplicado em seu corpo, mas agiam apenas por meio da presença do estímulo químico. As formigas não identificavam que havia algo de errado, considerando que apesar de suas irmãs estarem “exalando” a substância química relacionada à morte, elas se comportavam como se estivessem vivas. Segundo o teste dos zumbis, essas formigas seriam consideradas organismos destituídos de estados mentais e, conseqüentemente, animais sem consciência.

Contudo, isso não quer dizer que animais com cérebros simples não possam exibir comportamentos flexíveis. Já foi observado, por exemplo, que o nematoide *Pristionchus pacifus*, cujo cérebro possui cerca de 300 neurônios, apresenta comportamentos do tipo tomada de decisão. Esse organismo onívoro geralmente se alimenta de bactérias e larvas de outras espécies de nematoides, como as da *Caenorhabditis elegans*. Em um experimento recente, pesquisadores observaram que dependendo do ambiente em que a *P. pacifus* se encontra, ela altera os seus comportamentos de forrageamento face aos tipos de estímulos expostos. Foi visto que com uma única mordida a *P. pacifus* conseguia devorar uma larva de *C. elegans*, ao passo que a *C. elegans* adulta geralmente sobrevivia à mordida. O que parecia ser um comportamento predatório malsucedido da *P. pacifus* frente a *C. elegans* adulta, demonstrou ser um comportamento de proteção territorial. Em um ambiente com escassez de bactérias, a *P. pacifus* adotava uma estratégia de forrageamento predatório, cujo objetivo era morder e assim devorar as larvas de *C. elegans* que eram identificadas como presas. Enquanto no ambiente com abundância de bactérias, a *P. pacifus* mordida as *C. elegans* adultas com o objetivo de proteger o seu território de forrageamento, e não com o objetivo de as preda. Foi concluído que, em ambientes com abundância de bactérias, as mordidas da *P. pacifus* eram menos frequentes e geralmente tinham a função de proteção do ambiente com abundância de alimentos (bactérias). Por outro lado, em ambientes com escassez, as mordidas da *P. pacifus* eram mais frequentes e

tenham a função de predação. Os pesquisadores sugeriram que a depender do contexto, as *P. pacifus* consideravam os custos e benefícios de exibir um comportamento de mordida, motivando diferentes estratégias comportamentais. Em um ambiente com abundância de alimentos, era energeticamente mais vantajoso para a *P. pacifus* manter o seu território de forrageamento e morder as suas concorrentes com o intuito de afastá-las. Enquanto em ambientes com escassez, a estratégia mais vantajosa para a *P. pacifus* era gastar energia mordendo e devorando as larvas de seu concorrente nematoide (QUACH e CHALASANI, 2022). Considerando o que foi descoberto com o experimento, as *P. pacifus* passariam no teste dos zumbis?¹⁶⁰ Se sim, mesmo animais com sistemas nervosos extremamente simples poderiam passar no teste, o que para muitos soaria como algo negativo, uma vez que esvaziaria o papel que os estados mentais possuiriam na explicação de certos tipos de comportamentos.

Isto posto, apresentarei nas próximas duas subseções dois argumentos contrários à abordagem newtoniana de Tye. O primeiro é o denominado “Argumento da Inevitabilidade da Teoria” proposto por Birch (2020a) e o segundo é o que chamarei de “O Desafio de Carruthers-Dawkins”. Vejamos ambos a seguir.

3.5.1 O Argumento da Inevitabilidade da Teoria

Na maior parte de seu livro de 2017, Tye discorre sobre possíveis invalidadores que poderiam comprometer a Regra de Newton de diferentes formas, à luz dos mais diversos casos, indo dos peixes aos artrópodes. Com efeito, a pergunta que agora devemos fazer é a seguinte: o que seria um invalidador para Tye? Alguns pesquisadores poderiam argumentar que a ausência de um córtex, por exemplo, constituiria um invalidador (KEY, 2015, 2016). Tye defende que a ausência do córtex não constitui um invalidador. O argumento de Tye (2017) se baseia em evidências encontradas em crianças com um tipo específico de decorticação congênita, a hidranencefalia (Figura 13). Esta é uma condição rara onde há destruição severa da região cortical bilateral. Parte do córtex destruído é substituído pelo líquido cefalorraquidiano. O diagnóstico pode ser feito ainda durante a gestação utilizando-se, por exemplo, a Tomografia Computadorizada (TC), o EEG e a Ressonância Magnética Fetal (RM) (DUARTE et al, 2019; KANDEL et al, 2021, p. 981). Em geral, os bebês que nascem com essa condição falecem com poucas semanas de vida.

¹⁶⁰ Também já foi observado a presença de comportamentos, mesmo que rudimentares, de *trade-off* em anêmonas-do-mar da Nova Zelândia (*Anthopleura aureoradiata*), organismos que sequer possuem um sistema nervoso central (HAAG e DYSON, 2014). Note que comportamentos *trade-off* são geralmente descritos como flexíveis na literatura.

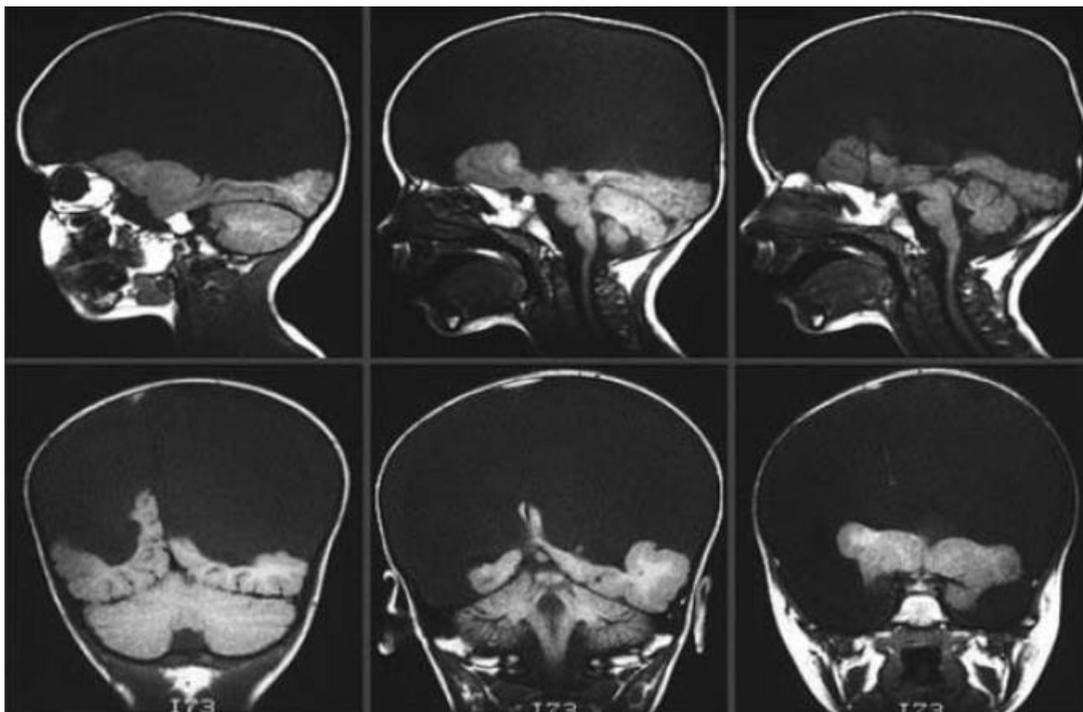


Figura 13 – Imagens oriundas de ressonância magnética. Plano sagital e frontal da cabeça de uma criança com hidranencefalia. Merker, 2007, p. 78.

Como Birch (2020a) bem observa, Tye lança mão da mesma evidência que o neurocientista Bjorn Merker (2005, 2007) utiliza para defender a sua teoria da natureza da consciência, a chamada “teoria do mesencéfalo”. Em linhas gerais, o mesencéfalo é uma estrutura subcortical evolutivamente mais antiga que o córtex cujos mecanismos são responsáveis por integrar *inputs* de diferentes fontes, orientando, assim, o organismo em seu ambiente (BARRON e KLEIN, 2016; MERKER, 2005, 2007). De acordo com tais evidências, as crianças hidranencéfalas supostamente seriam capazes de experimentar dor e outros estados mentais conscientes. Essas evidências são geralmente de natureza comportamental, como uma expressão que denotaria dor. Interessante notar que Merker (2007, p. 79) utilizou como evidências e-mails com as descrições comportamentais das crianças que foram enviados pelos seus cuidadores e contatos presenciais que duraram cerca de 1 semana com as crianças e as suas famílias. Segundo Tye (2017), as evidências encontradas em crianças com hidranencefalia demonstraria que a ausência de um córtex não constituiria um invalidador, uma vez que embora não possuam boa parte do córtex essas crianças são conscientes. O princípio tácito dos invalidadores de Tye seria, na interpretação de Birch (2020a), o seguinte:

*[...] para mostrarmos que a ausência de alguma estrutura neuroanatômica (como o córtex) não é um **invalidador (defeater)**, devemos encontrar*

*evidências em humanos mostrando que cérebros sem tal estrutura ainda possuem experiência consciente.*¹⁶¹ (BIRCH, 2020a, p. 7)

Segundo Birch (2020a), contudo, esse princípio acaba por suprimir a vantagem que a abordagem neutra possuiria sobre a teórica. A razão é a seguinte: a implementação do princípio supracitado acaba por se assentar em uma determinada teoria da consciência humana. Isso se dá na medida em que Tye recorre a uma evidência que não somente fornece suporte, mas faz parte da constituição central da teoria do mesencéfalo de Merker (2005, 2007). Aceitar a evidência de Merker implica em traçar condições mínimas para que haja consciência. Ao comprar tal evidência Tye se compromete, ao menos em parte, com uma tese específica da base neuronal e/ou da própria natureza da consciência. Sendo justo, mais recentemente Tye (2021, p. 105) afirmou que em indivíduos neurotípicos o córtex pré-frontal é de fato a base neuronal da consciência. Contudo, em casos de indivíduos neurodivergentes essa base poderia mudar em razão de processos internos envolvendo a capacidade de neuroplasticidade. No caso específico das crianças hidranencéfalas, Tye (2021, p. 107) afirma da poltrona que as populações neuronais de estruturas subcorticais poderiam sustentar os seus estados conscientes, mesmo que esses estados fossem mais rudimentares que os dos indivíduos neurotípicos. Tye não explica que tipo de processos neuroplásticos são esses, tampouco explica como tais processos funcionariam para que a consciência fosse transferida de estruturas neuronais tão diferentes entre si. A estratégia de Tye soa completamente injustificada e *ad hoc*, dado que ele não apresenta argumentos ou evidências empíricas que sustentem tal hipótese. O único motivo para a sua invocação seria o de garantir a realizabilidade da consciência em estruturas cerebrais mais simples e primeiras que as do córtex pré-frontal, possibilitando que humanos e animais sem córtex sejam capazes de realizar estados mentais conscientes. Ao menos da forma como é colocado por Tye, não vejo motivos para que se leve tal afirmação a sério.

Em suma, de acordo com Tye (2017), para que se verifique se a ausência do córtex constituiria ou não um invalidador, deve-se recorrer a evidências de humanos que não possuem um córtex, mas que ainda assim sejam conscientes. Deste modo, a evidência que for encontrada servirá como uma forma de condição suficiente e mínima para a realização da consciência. O descobrimento de casos reais de humanos conscientes sem um córtex tornaria convincente uma teoria da consciência motivada por uma estrutura subcortical como o mesencéfalo (BIRCH,

¹⁶¹ [...] *to show that the absence of some neuroanatomical structure (such as the cortex) is not a **defeater**, we need to find evidence in humans showing that brains without that structure still have conscious experience.*

2020a). Neste caso, o córtex seria considerado uma estrutura dispensável para a existência da consciência, mesmo que ele desempenhe um papel cognitivamente enriquecedor, ao passo que o mesencéfalo seria a estrutura minimamente suficiente para a realização da consciência.

Portanto, quando a Regra de Newton recorre a essa evidência, ela perde a sua pretensa neutralidade, uma vez que acaba se comprometendo tacitamente com uma teoria da consciência. Em outras palavras, ao utilizar tal evidência como suporte, a Regra de Newton perde a sua neutralidade no debate da literatura sobre a consciência, dado que a evidência apresentada garante a própria constituição de uma teoria específica da consciência, a saber, a teoria do mesencéfalo. Além disso, as evidências fornecidas por Merker (2005) são consideradas extremamente controversas dentro da literatura sobre o tema.

Os autores que defendem que o córtex pré-frontal, o córtex V4 ou qualquer outra região cerebral, que não o mesencéfalo, é a base neuronal da consciência, rejeitariam categoricamente a evidência apresentada por Merker e endossada por Tye. Note que considerar o córtex pré-frontal, por exemplo, como a região responsável pela consciência em humanos, não implica no compromisso com a seguinte afirmação forte: para que um animal *S* seja considerado consciente, ele deve possuir exatamente a mesma estrutura cerebral que em humanos é responsável pela consciência. Segundo essa afirmação, *S* só seria considerado consciente se tivesse um córtex pré-frontal. No entanto, muitos autores ficariam satisfeitos com uma versão moderada dessa tese segundo a qual os animais só precisariam ter uma região cerebral que *funcionasse analogamente* ao córtex pré-frontal dos humanos. Assim, o animal em questão só necessitaria ter uma região *funcionalmente análoga* à área responsável pela consciência em humanos. De acordo com essa interpretação mais moderada, uma ave que possuísse uma estrutura anatômica como o nidopálio caudolateral poderia ser considerado consciente, uma vez que essa estrutura funcionaria de forma análoga ao córtex pré-frontal dos mamíferos (GÜNTÜRKÜN, 2005).

Ademais, a evidência de crianças hidranencéfalas enfrenta uma série de problemas. Por exemplo, não é explicado satisfatoriamente por Merker como é possível a existência de um fenômeno como a visão cega (*blindsight*) (Figura 14) dado o contexto de seu *framework*. A visão cega é uma condição em que os indivíduos que tiveram uma lesão no córtex V1 possuem a negligência de uma área do seu campo visual. Eles relatam não possuírem consciência dos eventos que ocorrem nesta região. Todavia, o desempenho desses indivíduos em tarefas que dependam da percepção dos eventos que ocorrem no seu campo visual negligenciado se equipara com os de indivíduos neurotípicos (CHURCHLAND, 1986, pp. 224-228;

WEISKRANTZ, 1986). As informações perceptuais chegam normalmente no mesencéfalo intacto dos indivíduos com visão cega, porém não chegam em seu córtex V1. Ou seja, embora o mesencéfalo esteja intacto, os indivíduos com visão cega não são capazes de perceber conscientemente os estímulos visuais exibidos no seu campo visual negligenciado. Ora, isso é contraditório, pois como o mesencéfalo desses indivíduos se encontra intacto, e tal estrutura cerebral é responsável pela realização da consciência, seria de se esperar que mesmo com a lesão no córtex V1 os indivíduos permanecessem realizando estados perceptuais conscientes. Me parece que neste caso a consciência perceptual visual do sujeito dependeria em certa medida do córtex V1, e não do mesencéfalo. Essa objeção, portanto, vai de encontro à suposta evidência do mesencéfalo como condição necessária ou suficiente para a realização da consciência (BIRCH, 2020a; DOESBURG e WARD, 2007; SCHLAG, 2007; WATKINS e REES, 2007).

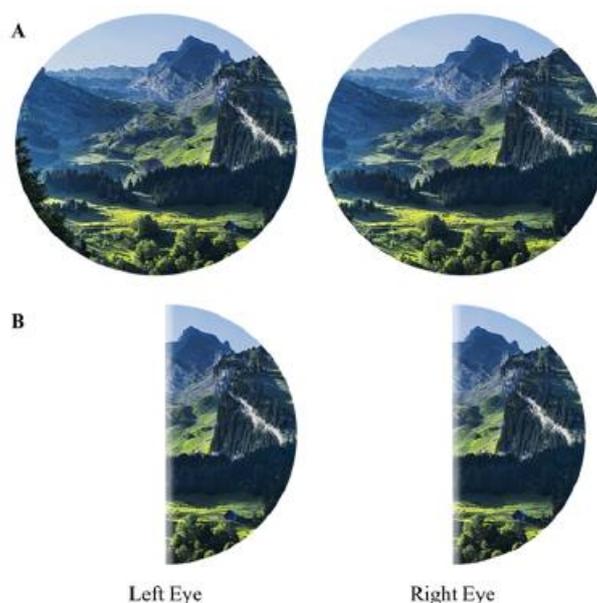


Figura 14 – A) Paisagem montanhosa percebida por um indivíduo neurotípico. B) Paisagem montanhosa percebida por um indivíduo com visão cega ocasionada por uma lesão no hemisfério direito do cérebro. Campo visual cego aproximado. Derrien et al, 2022, p. 2.

Mesmo diante de casos mais controversos, como os das abelhas, Tye não crê que a ausência do córtex, do mesencéfalo, ou outras diferenças radicais em sua neuroanatomia constituam um invalidador. Tye (2017, p. 150) sustenta que é devido à funcionalidade análoga entre a neuroanatomia e a arquitetura cognitiva das abelhas e dos humanos que a ausência das regiões cerebrais supracitadas nas abelhas não constitui um invalidador. Tye (2017, pp. 150-153) utiliza como ilustração as partes dos cérebros das abelhas conhecidos como *mushroom*

bodies. Segundo Tye (2017, pp. 150-153), essas estruturas seriam funcionalmente análogas ao córtex humano, garantindo a aplicação da Regra de Newton.¹⁶²

Em suma, Birch (2020a) formula o problema nos seguintes termos. Primeiro, para eximir a Regra de Newton da acusação de credulidade, Tye apresenta os seus invalidadores. No entanto, para sabermos se uma determinada evidência *P* constitui ou não um invalidador, exige-se a avaliação sobre o que a ausência de *P* implicaria para o caso dos humanos. Em outras palavras, para não cair em credulidade a Regra de Newton necessita de invalidadores que funcionam como “requisitos mínimos” para que a consciência seja realizada em um certo sistema cognitivo (BIRCH, 2020a). Entretanto, a Regra de Newton acaba por se comprometer tacitamente com alguma teoria T_n que diz respeito à consciência humana. Ou seja, se torna *inevitável* a apelação à T_n . Birch (2020a, p. 8) considera que o empreendimento de Tye é ainda mais ambicioso e difícil que os das teorias da consciência, pois a Regra de Newton não requer somente que a teoria diga respeito a condições suficientes e a marcadores positivos que indiquem a presença da consciência, mas também aos requisitos mínimos e/ou marcadores negativos. Deste modo, os defensores da abordagem teórica neutra nunca teriam à disposição evidências incontestáveis aos olhos de uma outra teoria concorrente, a não ser àquela subjacente à própria teoria (BIRCH 2020a, p. 8). Assim, a Regra de Newton fracassaria em resolver o debate sobre os casos de animais.

3.5.2 O Desafio de Carruthers-Dawkins

Ambos Peter Carruthers (2018a, 2019) e Marian Dawkins¹⁶³ (2017, 2021) são céticos quanto à viabilidade da abordagem neutra atribuir consciência aos animais. Segundo esta posição, em razão da Regra de Newton se assentar em uma tese que identifica a consciência com todo tipo de percepção, a chamada “Visão Simples”, ela negligenciaria vários modos de comportamentos flexíveis sensíveis ao ambiente, que são causadas por diferentes formas de estados perceptuais *inconscientes*. A Visão Simples acarretaria que toda forma de estado sensorio-perceptivo, como a percepção visual, garantiria a aplicação da Regra de Newton

¹⁶² Outros autores seguem uma linha parecida ao aplicar a teoria do mesencéfalo aos insetos (MERKER, 2005, 2007). Barron e Klein (2016) são um exemplo. Eles defendem que o complexo central (CX) dos insetos seria funcionalmente análogo ao mesencéfalo humano. Segundo eles, seria razoável sustentar que os insetos possuem consciência, dado que eles possuem uma estrutura funcionalmente análoga ao mesencéfalo (que em humanos é a estrutura responsável pela realização da consciência). Para maiores informações sobre o CX, cf. Evans-Turner e Jayaraman, 2016.

¹⁶³ Embora a crítica de Dawkins não tenha Tye como o alvo direto, por serem mais gerais, as suas reticências atingem todos os tipos de abordagens teóricas neutras. Deste modo, creio que podemos direcionar as suas críticas à posição de Tye.

(CARRUTHERS, 2019). Isto implicaria na incapacidade de aplicarmos a Regra de Newton nos animais, pois ela não faria uma distinção fina entre os estados perceptuais conscientes e inconscientes, e como cada um atua sobre os comportamentos de cada animal. Chamarei este tipo de argumento de “O Desafio de Carruthers-Dawkins”. Vejamos como ele funciona.

O mote central do Desafio de Carruthers-Dawkins é o seguinte: a Regra de Newton está embasada em uma tese que identifica a consciência com todo tipo de estado perceptual e afetivo (*affective*), caso se queira realizar uma separação entre os estados afetivos (também chamados de emotivos) e os exteroceptivos/perceptuais (*exteroceptive*) (FEINBERG e MALLATT, 2018; GODFREY-SMITH, 2020b).¹⁶⁴ No entanto, ela desconsideraria os estados perceptuais *inconscientes* que geram comportamentos flexíveis face aos estímulos do ambiente. Deste modo, diante de casos em que o comportamento engajado fosse flexível, mas que a sua causa fosse uma percepção inconsciente, não poderíamos aplicar a Regra de Newton. Para que pudéssemos aplicar a Regra de Newton seria fundamental a realização de uma distinção fina entre percepções conscientes e inconscientes. Porém, tal distinção exige uma teoria robusta da consciência.¹⁶⁵ Por outros meios, estamos de volta com a acusação do “Argumento da Inevitabilidade da Teoria”, apresentado na subseção anterior (BIRCH, 2020a). O sucesso da Regra de Newton depende de uma teoria substantiva da consciência com base na qual distinguiríamos os estados perceptuais conscientes dos inconscientes. Todavia, se a Regra de Newton o fizer, fracassará enquanto um princípio que pretende ser neutro. A Regra de Newton, portanto, seria um projeto natimorto. Para tornar mais claro, vejamos os seguintes exemplos.

Considere a seguinte ilustração fornecida por Carruthers (2018a, 2019). Imagine um confronto entre um arremessador e um rebatedor durante um jogo de beisebol. O arremessador lança uma bola rápida (*fast ball*) que viaja a 96 milhas por hora. A distância entre a base do arremessador e a do rebatedor é de cerca de 18 metros, e a distância entre o ponto da qual o arremessador lança a bola e o ponto de contato dela com o bastão do rebatedor é de 16 metros. A bola arremessada demorará cerca de 400 milissegundos para percorrer tal distância. Durante este curto espaço de tempo, o rebatedor terá que estimar a rota e o tipo de arremesso que será realizado, iniciar o swing, executar o tipo de ação motora que o dado arremesso demanda para, assim, conseguir um bom contato com a bola. No entanto, a nossa melhor

¹⁶⁴ Não me comprometerei com nenhuma tese acerca da natureza da consciência. Assim, me manterei neutro em relação ao debate se todo estado consciente é inteiramente perceptual ou não.

¹⁶⁵ “[...] *one cannot know which perceptual states are conscious and which of them cause any given type of behavior in the absence of a good theory of consciousness.*” (CARRUTHERS, 2019, p. 56) “*It is plain that Tye is mistaken. In order to apply Newton’s Principle correctly, we need a theory of the difference between conscious and unconscious forms of perception. Armed with such a theory, we can then raise (and investigate) the question whether both forms of perception are manifested in animals.*” (CARRUTHERS, 2018a, p. 186)

estimativa sobre o tempo que leva para o estímulo, a luz refletida pela bola, atingir a nossa retina e ser processado pelo sistema visual para se tornar consciente estaria em por volta de 350 milissegundos (DEHAENE, 2014). Como Carruthers (2018a, 2019) assinala, o swing do rebatedor seria iniciado antes da percepção do estímulo visual se tornar consciente. O comportamento do rebatedor nos parece flexível, dado a quantidade de situações que o jogador tem que levar em conta e a forma com que ele tem que ajustar o seu movimento de acordo com o tipo de arremesso executado. Este seria um caso de comportamento flexível causado por uma percepção *inconsciente*. De acordo com Carruthers (2019), o comportamento do rebatedor seria *acompanhado* por um estado consciente, mas não *causado* por ele. Qual o tipo de conclusão que podemos retirar deste exemplo? A impossibilidade da aplicação da Regra de Newton no caso dos animais. Segundo Carruthers (2018a, 2019), quando observamos um animal realizando um tipo de comportamento similar a este, em resposta a um estímulo similar fornecido pelo ambiente, não seríamos capazes de aplicar a Regra de Newton. O motivo seria o de que não conseguiríamos saber se o comportamento do animal teria sido *causado* ou apenas *acompanhado* por um estado mental consciente (caso ele fosse capaz de realizar algum).

Um segundo exemplo emerge da descoberta dos neurocientistas David Milner e Melvyn Goodale. Ela diz respeito à existência de dois tipos de processos nos sistemas visuais dos humanos. São eles o fluxo ventral e o fluxo dorsal, cada um atravessando uma determinada rota ao longo do córtex (Figura 15) (GOODALE, 2014; GOODALE e MILNER, 2017; MILNER e GOODALE, 2006). O fluxo ventral segue do lobo occipital, especificamente no córtex visual, à região inferior do lobo temporal. Já o fluxo dorsal sai da mesma região posterior do córtex, mas vai em direção ao lobo parietal. O fluxo ventral inclui em sua formação as regiões V1, V2, V3, V4 e o córtex infratemporal. O fluxo dorsal, por sua vez, vai do córtex V1, V2, V3, córtex medial temporal até o córtex temporal medial superior (EYSENCK e KEANE, 2020, p. 47). Claro, no início do processamento, os dois fluxos se interconectam, apesar de depois se separarem. Como função, o fluxo ventral realiza a construção das representações visuais do ambiente, resultando no processamento de reconhecimento dos objetos e eventos e, assim, atribuindo significado e estabelecendo relações causais. Ao passo que o fluxo dorsal é responsável por guiar as nossas ações em tempo real e por processar as informações da disposição espacial dos objetos (GOODALE, 2014). Em relação ao aspecto evolutivo, alguns autores sugerem que esses dois tipos de processamento de informações visuais contribuíram para a visão noturna apurada dos primatas ancestrais se comparados com outros mamíferos de seu tempo. Essa adaptação teria posteriormente facilitado os humanos a reconhecerem amplos

conjuntos de indivíduos e, assim, formarem grandes grupos sociais. Isso, por sua vez, propiciou a capacidade de elaboração de ferramentas, no surgimento da linguagem e outras atividades cognitivamente complexas (KAAS et al, 2021).

Uma das formas que se verificou esses dois modos de processamento visual foi através de pacientes com lesões neurológicas. Os pacientes que tiveram danos no córtex parietal posterior e na região do sulco intraparietal, partes integrantes do *fluxo dorsal*, padecem de *ataxia ótica*. Nesta condição, o paciente possui dificuldade de executar movimentos precisos guiados pela visão, resultando em comportamentos desordenados. Todavia, a sua experiência consciente permanece intacta (CARRUTHERS, 2019; EYSENCK e KEANE, 2020, p. 57; GOODALE, 2014). Por outro lado, os pacientes com lesões nas regiões ventrolateral do lobo occipital, integrantes do *fluxo ventral*, padeceriam de *agnosia visual*. Esta condição torna o sujeito incapaz de identificar objetos e formas de modo consciente, porém mantendo intactos os seus movimentos motores (EYSENCK e KEANE, 2020, p. 58; GOODALE, 2014; JAMES et al, 2003).

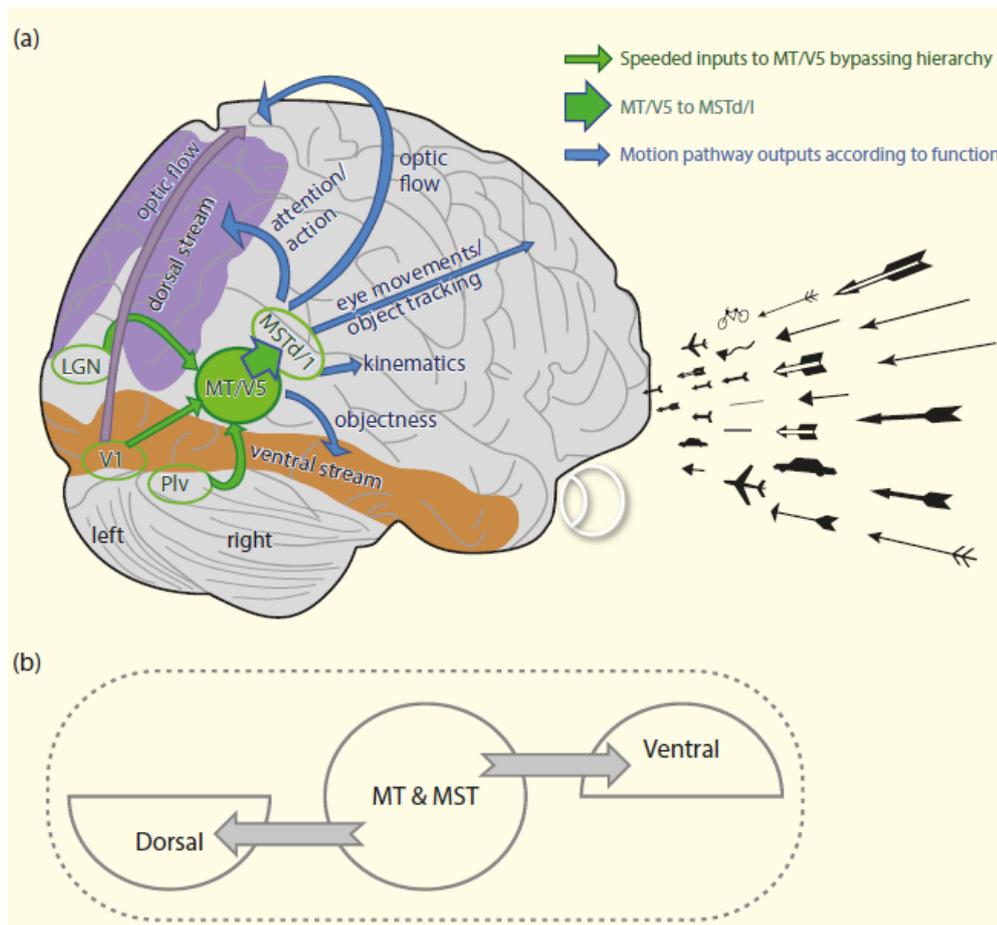


Figura 15 – Fluxo ventral e dorsal. Eysenck e Keane, 2020, p. 53.

Considere o clássico caso da paciente D. F, uma jovem mulher que sofreu um dano cerebral irreversível. Devido a uma hipoxia através de envenenamento por monóxido de carbono, D. F. teve uma série de lesões em seu cérebro, notadamente no córtex bilateral ventral lateral-occipital (JAMES et al, 2003). Como resultado, D. F. teve um dano no seu fluxo ventral, ocasionando uma severa agnosia visual de forma (*visual-form agnosia*). Apesar de experienciar as cores e texturas dos objetos, a paciente não conseguia experienciar as formas que eles possuíam. Ao pegar em uma banana, por exemplo, ela não conseguia identificar a fruta somente por meio do contato com as curvas da sua forma. D. F. não conseguia nem mesmo distinguir formas geométricas simples, como um bloco de madeira quadrado de um retangular. No entanto, a sua coordenação motora permaneceu intacta. A sua capacidade de guiar as mãos ao longo do objeto e pegá-lo era indistinguível de uma pessoa neurotípica. D. F. também era capaz de colocar cartas em fendas de caixas disposta em ângulos diferentes de forma similar ao de um indivíduo neurotípico. (GOODALE, 2014; JAMES et al, 2003; MILNER et al, 1991). Em um outro experimento, os pesquisadores utilizaram 6 pares de formas diferentes para verificar a capacidade de discriminação e de agarrar de D. F. Ao mostrarem pares de formas iguais ou diferentes, D. F. não era capaz de discriminar se as formas do par eram iguais ou não. Contudo, quando pedido para que ela pegasse com precisão essas formas utilizando o dedo indicador e o polegar, D. F. desempenhava tal tarefa de maneira similar ao de um indivíduo neurotípico (GOODALE et al, 1994).

Outro exemplo é o de pessoas com visão cega (Figura 14). Como visto anteriormente, essa é uma condição na qual os pacientes que tiveram lesão no córtex V1 ficam totalmente ou parcialmente privadas de consciência em uma determinada porção do seu campo visual (o chamado “escotoma”) (HARLEY, 2021; WEISKRANTZ, 1986). No entanto, elas ainda conseguem identificar formas e padrões de movimento dos objetos apresentados no campo visual cego. Quando apresentadas a um objeto e perguntadas se as viam, os pacientes respondiam negativamente, porém quando exortadas eles respondiam com grande grau de acerto semelhante às pessoas neurotípicas (DEHAENE, 2014). Ou seja, os pacientes com visão cega exibiam comportamentos que visto por fora, eram parecidos com o de uma pessoa neurotípica. Distinguir somente por meio do comportamento uma pessoa com visão cega de uma neurotípica se tornaria uma tarefa difícil. O paciente P. S., por exemplo, possuía uma negligência na porção esquerda do seu campo visual. Os pesquisadores fizeram o seguinte experimento com ele. Mostravam-lhe uma imagem que continha uma figura de uma casa pegando fogo na parte esquerda e uma casa normal na parte direita e o perguntavam em qual

delas ele gostaria de morar. Apesar de ter relatado não ter percebido conscientemente diferenças entre as duas casas, em sua resposta ele sempre escolhia a casa da direita (MARSHALL e HALLIGAN, 1988). Somado a isso, existem evidências de que macacos, animais filogeneticamente próximos de nós, com visão cega conseguiriam realizar até mesmo formas de aprendizado instrumental visual¹⁶⁶ (KATO et al, 2021). Isto é, associar o seu comportamento, neste caso uma resposta a um estímulo visual, com as consequências resultantes dele (recompensa ou punição) (DOMJAN, 2014).¹⁶⁷

Por fim, o último exemplo que fornecerei será de natureza afetiva. Os estados afetivos, ou emocionais, como os de medo, raiva e prazer geralmente provocam uma experiência consciente nos humanos, resultando em comportamentos diferentes e flexíveis. Todavia, aquilo que chamamos de “medo” consiste em respostas comportamentais com circuitos e processamentos neuronais distintos. O mesmo estado de medo que provoca comportamentos instintivos, autônomos e flexíveis como *fight-or-flee* (lute ou fuja), aumento da taxa de batimento cardíaco e respiração ofegante, também ocasiona o estado consciente do medo. O estado mental de medo, por diferentes vias de processamentos neuronais, pode ocasionar comportamentos que são meramente reflexos inatos e comportamentos flexíveis que são *inconscientes*, e outros comportamentos flexíveis que são gerados pelo estado mental *consciente* do medo (CARRUTHERS, 2019; DAWKINS, 2017, 2021; LEDOUX, 2014). Por exemplo, K.-H. J. era um indivíduo de 65 anos que sofreu uma série de infartos em ambos os lados da artéria cerebral posterior ocasionando uma completa cegueira cortical bilateral (HAMM et al, 2003). O indivíduo foi diagnosticado com visão cega e agnosia. Ele não conseguia pegar objetos em seu campo de visão e nem orientar a sua visão em resposta a novos estímulos visuais. No entanto, os pesquisadores conseguiram realizar um condicionamento no indivíduo, de modo em que ele associou inconscientemente um estímulo visual com um certo evento. Quando o mostravam um sinal visual que era seguido por um evento aversivo, um choque elétrico desconfortável sem gerar dor, o indivíduo apresentava comportamentos tipicamente associados a medo, como piscar o olho como forma de proteção, resposta galvânica da pele e aceleração da frequência de batimentos cardíacos (HAMM et al, 2003; LEDOUX, 1998). Ao perguntarem se tinha visto os sinais visuais, K.-H. J. respondia que não tinha visto

¹⁶⁶ Aqui utilizo os termos “operante” e “instrumental” de modo intercambiável. Cf. Allen, 2018.

¹⁶⁷ Em homenagem a um dos prováveis descobridores do condicionamento operante, o psicólogo B. F. Skinner, Daniel Dennett (1995, 2017) denominou de “criaturas skinnereanas” aquelas que apresentam esta forma de aprendizado. Cf. também, Godfrey-Smith, 2018a.

nada, pois estava cego. Ora, isso demonstra que ele percebeu inconscientemente o estímulo visual apresentado e realizou um estado afetivo inconscientemente.

Conjuntamente com Carruthers (2018a, 2019), considero que a tese na qual a Regra de Newton se assenta, a saber, a Visão Simples, está equivocada, uma vez que ela ignora os diversos casos de percepções *inconscientes* que geram comportamentos flexíveis. Não seria prudente identificar todos os casos de percepção como conscientes. A Regra de Newton demonstra, portanto, estar em um beco sem saída. Para que ela consiga distinguir os estados conscientes dos inconscientes ela necessita apelar a uma teoria da consciência, o que resulta no seu insucesso enquanto proposta teoricamente neutra. Caso ela negligencie esta distinção, a sua aplicação estaria também condenada ao fracasso, dado as diversas evidências de estados perceptuais e afetivos inconscientes que produzem comportamentos flexíveis. A Regra de Newton seria nada mais que um projeto natimorto.

No entanto, o fracasso da Regra de Newton não implicaria na condenação de todas as propostas epistêmicas. Como veremos na próxima seção, a abordagem teórica leve consegue se esquivar dos problemas elencados acima. Ademais, ela possui a vantagem de ser a hipótese epistêmica filosoficamente mais plausível e empiricamente adequada.

3.6 A Abordagem Teórica Leve

Como forma de contornar os equívocos da abordagem newtoniana de Tye (2017, 2018) e das abordagens teóricas, Birch (2020a) propõe uma abordagem que consiste em um meio termo entre elas, a chamada “abordagem teórica leve”. Essa abordagem não seria nem totalmente neutra, como a de Tye (2017), nem completamente comprometida com uma determinada teoria da consciência humana (T_n), como as abordagens de Carruthers (2019) e Prinz (2018). A abordagem teórica leve busca se esquivar do comprometimento prévio com T_n , ao mesmo tempo que apresentaria diminutos compromissos teóricos, o que não a torna neutra.

A abordagem teórica leve propõe uma hipótese “ampla” sobre a relação entre a consciência e a cognição dos animais e dos humanos. Birch (2020a, p. 8) a denomina de “*hipótese da facilitação*”. Segundo essa hipótese, dadas as leis da natureza em nosso mundo atual, a consciência facilitaria a cognição na execução de determinadas funções, habilidades e tarefas cognitivas.¹⁶⁸ Como o próprio Birch admite, ainda está em aberto qual seria exatamente

¹⁶⁸ Suspeito que a chamada “tese da complexidade ambiental” (TCA) de Godfrey-Smith (1998) antecipa, em alguma medida, a hipótese da facilitação de Birch. A TCA expressa que: “*The function of cognition is to enable the agent to deal with environmental complexity.*” (GODFREY-SMITH, 1998, p. 3). Nesse sentido, a cognição facilitaria o organismo a lidar com os problemas impostos pela complexidade de seu ambiente. Ora, a hipótese de Birch dá um passo atrás, na medida em que a própria consciência desempenharia um papel facilitador na cognição.

a contribuição que a consciência faz para a cognição. No entanto, buscarei mostrar nesta seção que existem diversas evidências na literatura que apontam a consciência como o facilitador da realização de determinadas funções cognitivas. De forma mais estrita, a hipótese da facilitação expressa que:

*A consciência fenomenal perceptual de um estímulo facilita, relativa à percepção inconsciente, um conjunto de habilidades cognitivas em relação a esse dado estímulo.*¹⁶⁹ (BIRCH, 2020a, p. 8)

Como pode-se observar, a hipótese da facilitação se assenta na distinção entre os processos perceptuais *conscientes* e *inconscientes* quando um sistema cognitivo percebe um dado estímulo. Ou seja, ela segue implicitamente uma metodologia de *contraste mínimo*, de modo similar à que os teóricos da GWT realizam, estabelecendo situações capazes de distinguir um estímulo percebido conscientemente do inconsciente (DEHAENE, 2014; MASHOUR et al, 2020). Claro, disso não se segue que a hipótese da facilitação esteja comprometida com a GWT, mas somente com um certo tipo de método de pesquisa que, porventura, a GWT também utiliza. Embora Birch não seja claro, aqui me parece que ele está tacitamente respondendo ao Desafio de Carruthers-Dawkins, visto que ele distingue percepções conscientes das inconscientes. Mais do que isso, Birch coloca como crucial tal distinção, o que lhe afasta da abordagem newtoniana. De fato, a vantagem que ele toma em fazer essa distinção desta maneira é não se comprometer com uma teoria específica da consciência humana, mas somente com o conjunto (*cluster*) de habilidades cognitivas e a sua relação com a consciência.

Dito isso, algumas clarificações devem ser feitas. Primeiro, de acordo com Birch devemos entender o termo “facilitação” como “facilita-se, quando tudo está sob ordem”¹⁷⁰ (BIRCH, 2020a, p. 8). Ou seja, quando as condições estão normais, a saber, o estímulo apresentado, a dificuldade da tarefa, o modo de execução entre outras, ao perceber conscientemente um estímulo *x* um dado conjunto de habilidade cognitivas são facilitadas pela consciência (BIRCH, 2020a, p. 8). Segundo, devemos entender o termo “conjunto” (*cluster*) de

Nesse sentido, a consciência tornaria mais fácil para o organismo a execução de determinadas tarefas cognitivas. Também inspirando-se na TCA, Veit (2022a) propôs a chamada “tese da complexidade patológica”. Ela expressa que: “*The function of consciousness is to enable the agent to respond to pathological complexity.*” (VEIT, 2022a, p. 1). Essa tese salienta a importância da consciência nas estratégias evolutivas que os organismos realizam durante a sua história de vida. Note que ambas as propostas de Birch e Veit sublinham a importância da consciência na cognição e no desempenho de certos comportamentos.

¹⁶⁹ “*Phenomenally conscious perception of a stimulus facilitates, relative to unconscious perception, a cluster of cognitive abilities in relation to that stimulus.*”

¹⁷⁰ “[...] *facilitates, holding all else fixed.*”

habilidade cognitivas como um conjunto de marcadores de consciência que está ligado putativamente à consciência. Os marcadores de consciência desse conjunto possuem uma relação de covariação, isto é, eles estão correlacionados, indo e vindo juntos na medida em que o determinado estímulo é percebido ou não conscientemente (BIRCH, 2020a, pp. 8-9). Com isso, Birch sugere que não existiria um único marcador de consciência que deve ser buscado para determinar se um animal é ou não consciente. Nem mesmo existiria uma espécie de “lista de mercado” de marcadores de consciência, onde marcaríamos um *check* a cada novo marcador descoberto no animal. Na verdade, existiria uma série de marcadores de consciência que conjuntamente se ligam à consciência. Suponha que os marcadores de consciência M_1 , M_2 , M_3 e M_4 formem o conjunto (*cluster*) C_m . Todos esses marcadores estão conjuntamente ligados em uma relação de covariação. Para que se indique a presença da consciência em um dado animal S não devemos buscar um marcador em específico, como “ M_1 ”, tampouco verificar se S satisfaz uma determinada lista de marcadores, como ($M_1, M_2, M_3, M_n...$). Para que se possa indicar de forma confiável se S é ou não consciente, deve-se observar se S apresenta ou não C_m , pois é C_m que garantirá uma maior robustez na justificação acerca da hipótese de que S é consciente. Isto vai de encontro ao uso de um único marcador específico ou em uma mera checagem de marcadores em uma lista (BIRCH, 2020a).

Como vimos anteriormente, uma das grandes falhas da AP é justamente se assentar em um marcador único. Apenas um marcador não justifica a indicação da presença da consciência em um animal. Do mesmo modo que a satisfação de uma lista de mercado composta por n marcadores também não a justifica. Ambas as abordagens, a padrão e a da lista de mercado, estão fortemente suscetíveis a alarmes falsos ou falsos negativos. Poderíamos facilmente encontrar um animal que apresenta o marcador específico M_n , mas que talvez não seja consciente (e. g. os ratos que possuem a sua medula espinhal desconectada do cérebro, mas que apresentam comportamentos de evasão aprendidos) (ALLEN et al, 2009; GRAU, 2017). Contra a tese da lista de marcadores, poder-se-ia citar os casos de visão cega, onde embora não percebam conscientemente os estímulos exibidos em seus campos visuais negligenciados, eles desempenham uma série de tarefas e comportamentos que poderiam servir de marcadores (WEISKRANTZ, 1986). Assim, o grau de confiança que as teses do marcador único e da lista de mercado fornecem à justificação da hipótese de que o animal x possui consciência é consideravelmente menor do que a da abordagem teórica leve.

Com efeito, a abordagem teórica leve se compromete teoricamente com a existência de uma “entidade” que é o conjunto C_m e um certo modo de relação causal, a facilitação

(BIRCH, 2020a, p. 13). Por conseguinte, ela não se mantém totalmente isenta, como a abordagem teórica neutra e newtoniana e não se compromete com T_n tal como a abordagem teórica. Deste modo, ela se esquivava dos problemas que ambas as abordagens sofrem. Embora a abordagem teórica leve seja neutra face às diversas teorias da consciência, Birch (2020a, p. 9) a considera compatível com boa parte das teorias mais relevantes. Birch afirma que a hipótese da facilitação seria compatível com, por exemplo, a teoria do mesencéfalo de Merker (2005, 2007), a teoria da memória visual frágil de curto prazo de Block (2007a, 2011a), os representacionismos de primeira ordem (DRETSKE, 1995; TYE, 1995) e os de ordem superior (CARRUTHERS, 2000). Além da teoria do espaço de trabalho global (DEHAENE, 2014; MASHOUR et al, 2020) e da teoria da informação integrada (TONONI e KOCH, 2015). Isto se dá devido à todas as teorias supracitadas assumirem que a consciência desempenha algum tipo de papel facilitador para as habilidades cognitivas. Claro, cada teoria divergirá quanto a quais habilidades cognitivas em específico a consciência facilita. Mas o ponto de que a consciência desempenha esse papel é compartilhado entre todas elas (BIRCH, 2020a, p. 9).

Como já dito, a abordagem teórica leve não se mantém totalmente neutra como propõe a abordagem de Tye (2017). Pois, além da abordagem teórica leve se comprometer com C_m , ela descarta, por exemplo, qualquer tese do tipo epifenomenalista, em especial o que versa sobre a consciência, como o epifenomenalismo cognitivo. O epifenomenalismo cognitivo, quando trata da percepção, sustenta “a visão de que a percepção consciente (em oposição à inconsciente) de um estímulo não possui consequências alguma para a cognição.”¹⁷¹ (BIRCH, 2020a, p. 9). Segundo essa visão, todas as formas de habilidades cognitivas desempenhadas durante a percepção consciente de um dado estímulo também podem igualmente ser desempenhadas pela percepção inconsciente do mesmo estímulo. Ora, isso iria de encontro ao que a hipótese da facilitação expressa. Não entrarei nos meandros do debate acerca da plausibilidade do epifenomenalismo. Porém, não me parece razoável defender quaisquer espécies de epifenomenalismos. O epifenomenalismo recai em uma série de problemas, como mais notadamente a negação do fecho causal do mundo físico. Ademais, basta que se verifique a ampla rejeição que tal tese recebe por parte da literatura filosófica e mesmo científica.¹⁷² Esta é uma tese de poltrona destituída de evidências e que não desempenha nenhum papel explanatório relevante para a filosofia e a ciência da consciência, o que a torna uma noção dispensável e muito provavelmente incorreta.

¹⁷¹ “[...] *the view that conscious (as opposed to unconscious) perception of a stimulus has no consequences whatsoever for cognition.*”

¹⁷² Cf. Robinson, 2019.

Como Birch (2020a, p. 10) bem sublinha, a hipótese da facilitação não é uma tese *a priori* feita da poltrona por meio da razão pura. Na verdade, ela é empiricamente motivada e fundamentada em uma série de evidências disponíveis. Birch sugere três possíveis candidatos a marcadores de consciência que, dentro da literatura científica, se demonstram razoáveis. Todos eles consistiriam em tarefas cognitivas facilitadas pela consciência. Ou seja, quando o dado estímulo é percebido conscientemente, ao invés de inconscientemente, as suas realizações são facilitadas pela consciência. São os candidatos: o condicionamento de rastreo, a aprendizagem reversa rápida e a aprendizagem intermodal. Ademais, acrescento mais um: o paradigma da dissociação dupla (*double dissociation paradigm*). Como forma de economia, descreverei apenas o primeiro e o último candidato.¹⁷³

O primeiro candidato apontado por Birch (2020a) é o chamado “*condicionamento de rastreo*” (*trace conditioning*) (Figura 16). O condicionamento de rastreo é uma forma de condicionamento clássico, ou pavloviano, “na qual cada um dos dois estímulos é separado por um intervalo de tempo.”¹⁷⁴ (BIRCH, 2020a, p. 10). Um exemplo deste tipo de condicionamento é o que utiliza o rastreo do piscar de olhos (*trace eyeblink*). Neste tipo de experimento é fornecido ao participante um som próximo ao seu ouvido. Um segundo após o fornecimento do som, é direcionado um jato de ar nos olhos do participante, ocasionando a resposta instintiva de piscar os olhos (CLARK e SQUIRE, 1998, 1999; CLARK et al, 2001, 2002). Foi descoberto que os indivíduos só associavam o primeiro estímulo, o som, com o segundo, o jato de ar, quando tinham consciência de ambos os estímulos e do intervalo de tempo entre a exibição de cada um deles (BIRCH, 2020a). Isso foi descoberto quando os pesquisadores colocaram os participantes executando tarefas paralelas durante a apresentação dos estímulos. As tarefas paralelas distraíam os participantes, fazendo com que eles não conseguissem associar os estímulos e o intervalo de tempo entre eles. Perguntados após terem executados essas tarefas, os sujeitos responderam que não tiveram ciência (*awareness*) dos estímulos, tampouco do intervalo de tempo entre eles. Por outro lado, quando não se tinha que realizar tarefas paralelas, os participantes conseguiam aprender a relacionar os estímulos e o intervalo temporal e, assim, fechar os olhos imediatamente após a exibição do sinal sonoro. Concluiu-se que a consciência de alguma forma desempenha um papel facilitador neste tipo de condicionamento, uma vez que quando tinham que executar tarefas paralelas, os sujeitos não conseguiam associar os estímulos

¹⁷³ Para saber sobre os outros dois candidatos, cf. Birch, 2020a; Mudrik et al, 2014; Palmer e Ramsey, 2012; Travers et al, 2018.

¹⁷⁴ “[...] *in which the two stimuli are separated by a temporal interval.*”

e os intervalos temporais entre eles, além de relatarem não terem obtido ciência (*awareness*) da exibição dos estímulos e de suas relações temporais.

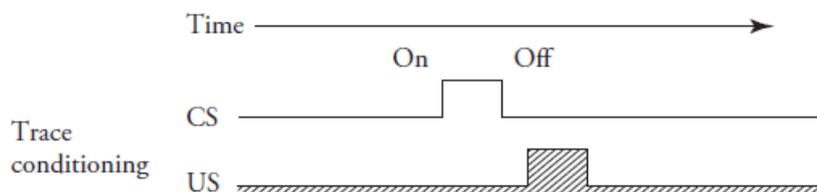


Figura 16 – Tradução da esquerda à direita. Tempo; Ligado; Desligado; Condicionamento de rastreamento; CS (Estímulo Condicionado); US (Estímulo Incondicionado). Domjan, 2014, p. 73. Modificado.

Por sua vez, em condições semelhantes, os indivíduos realizavam com sucesso o condicionamento com *delay* (*delay conditioning*) (Figura 17) executando ou não tarefas paralelas. O condicionamento com *delay* é aquele onde o estímulo condicionado (CS, sigla em inglês para *conditioned stimulus*) se inicia antes do estímulo incondicionado (US, sigla em inglês para *unconditioned stimulus*). O primeiro estímulo não dura o mesmo tempo que o segundo, mas existe uma sobreposição entre o primeiro e o segundo. Por exemplo, o som do sino (CS) se inicia e permanece durante um tempo enquanto a comida (US) é apresentada. O condicionamento com *delay* demonstra que existiriam processos mais simples que não demandariam a consciência para serem executados com sucesso. Em contraste com as evidências envolvidas no condicionamento de rastreamento que apontam que para a tarefa ser realizada com sucesso, a consciência deve estar presente facilitando as habilidades cognitivas correspondentes.

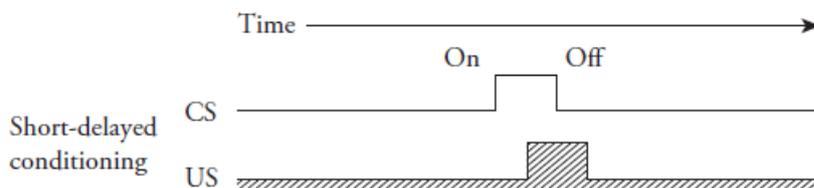


Figura 17 - Tradução da esquerda à direita. Tempo; Ligado; Desligado; Condicionamento com delay encurtado; CS (Estímulo Condicionado); US (Estímulo Incondicionado). Domjan, 2014, p. 73.

Em um artigo recente, Ben-Haim e colegas (2021) realizaram uma série de 7 experimentos com o objetivo de “*desentrelaçar*” os estados conscientes dos inconscientes. Para isso, eles utilizaram o paradigma da dissociação dupla modificado para atingir o objetivo. A ideia geral do paradigma é a seguinte: os pesquisadores apresentavam dois tipos de estímulos,

um *supraliminar* e outro *subliminar*. No primeiro, o indivíduo perceberia conscientemente o estímulo, pois ele seria apresentado de forma claramente visível. Enquanto no segundo estímulo, o indivíduo não perceberia conscientemente, pois ele seria apresentado rapidamente de modo a torná-lo invisível para o sujeito. Cada um desses estímulos gerava uma resposta comportamental diferente. Por exemplo, os participantes eram requisitados a completar uma palavra, como “SPI__”, sem utilizar uma palavra referência, como “SPICE”, que foi anteriormente apresentada de modo *supraliminar* ou *subliminar*. Os indivíduos expostos à referência *supraliminar* conseguiam corretamente responder uma outra palavra, como “SPIKE”. Por sua vez, quando o referente era apresentado de modo *subliminar*, por exemplo, “SPICE”, os participantes forneciam a palavra referência mais frequentemente do que o grupo controle, onde não era apresentado uma palavra referência. Em geral, os pesquisadores argumentam que aqui estão envolvidos dois tipos diferentes de processamentos perceptuais, a saber, um consciente e outro inconsciente. Como previsto, cada um desses processamentos perceptuais produziria uma forma distinta de resposta comportamental (DEHAENE, 2014). Colocando de forma geral, em humanos o processamento perceptual consciente P_c produziria uma resposta comportamental C_c , ao passo que o processamento perceptual inconsciente P_i produziria uma resposta comportamental C_i .

Apesar do experimento anterior fazer uso da linguagem, o paradigma geral não o necessita. Ben-Haim e colegas (2021) realizaram modificações no experimento de modo a torná-lo não verbal.¹⁷⁵ A tarefa sugerida no experimento era a seguinte: os participantes tinham que localizar um estímulo alvo em um dos dois lados da tela. No entanto, antes da apresentação do estímulo alvo, fornecia-se um sinal (*cue*) aos participantes. Este sinal (*cue*) se encontraria no lado oposto ao que posteriormente o estímulo alvo seria exibido. Ou seja, o sinal (*cue*) predizia o lado em que o estímulo alvo apareceria na tela. Por exemplo, se o sinal (*cue*) surgisse no lado direito da tela, necessariamente o estímulo alvo apareceria no lado esquerdo e vice-versa. Esse sinal (*cue*) era denominado de “estímulo incongruente”. O estímulo incongruente poderia ser apresentado de forma *supraliminar* ou *subliminar* ao participante (Figura 18).

¹⁷⁵ O n amostral foi de 4 macacos rhesus adultos e 145 humanos (BEN-HAIM et al, 2021, p. 8).

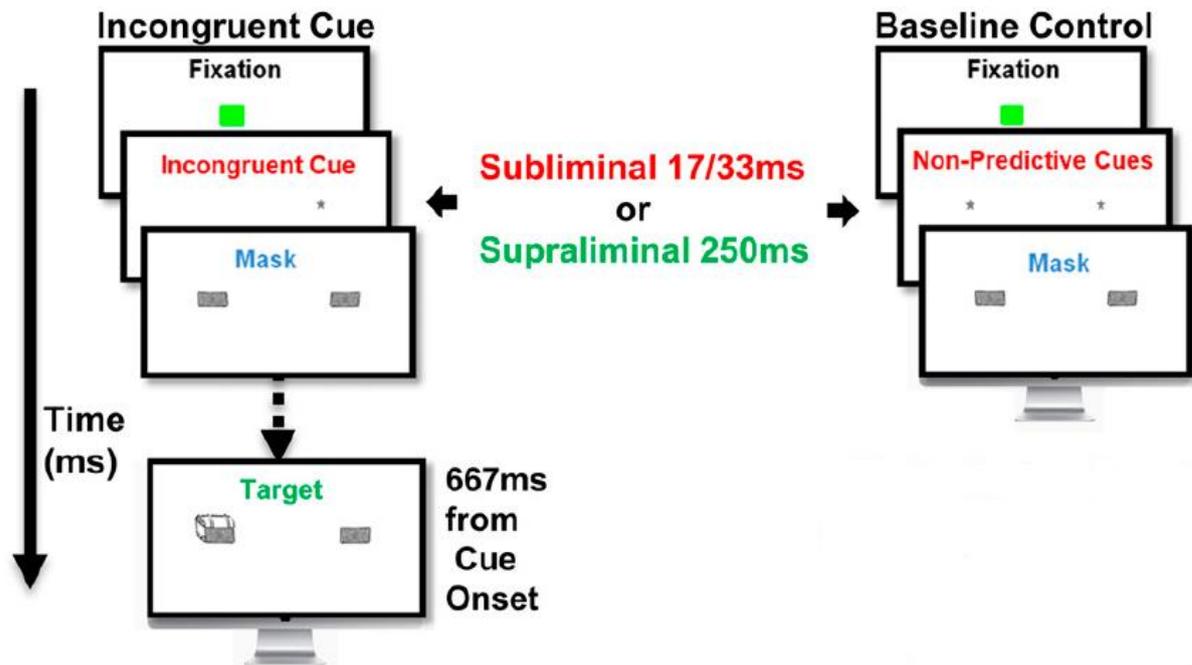


Figura 18 – Ben-Haim et al, 2021, p. 3. Modificado.

Foi observado que quando os participantes eram apresentados ao sinal (*cue*) de forma *supraliminar*, eles obtinham ciência de que o estímulo alvo estaria no lado oposto ao do sinal (*cue*) e assim eram mais rápidos no direcionamento da visão no estímulo alvo. Por outro lado, quando o sinal (*cue*) era fornecido de modo *subliminar*, os participantes eram mais lentos no direcionamento de sua visão no estímulo alvo. Isso se dava pois o sinal (*cue*) inconsciente atraía a atenção dos participantes para a sua própria localização. A consciência perceptual do sinal (*cue*), deste modo, facilitava o aprendizado dos participantes, fazendo com que eles relacionassem rapidamente o sinal (*cue*) com a localização do estímulo alvo. Enquanto a percepção inconsciente do sinal (*cue*) gerava uma resposta mais lenta dos participantes, na medida em que não havia facilitação para o aprendizado entre a relação do sinal (*cue*) com a localização do estímulo alvo.

Da mesma forma que os participantes humanos, foi verificado que os macacos rhesus (*Macaca mulatta*) também conseguiam identificar mais rapidamente a localização do estímulo alvo quando o sinal (*cue*) era fornecido de modo supraliminar. A cada acerto na localização do estímulo alvo, os macacos recebiam um suco como recompensa, motivando-os a executarem a tarefa com sucesso. Por sua vez, quando o estímulo incongruente era apresentado de forma subliminar, os macacos demoravam mais tempo para localizar o estímulo alvo projetado na tela (Figura 19).

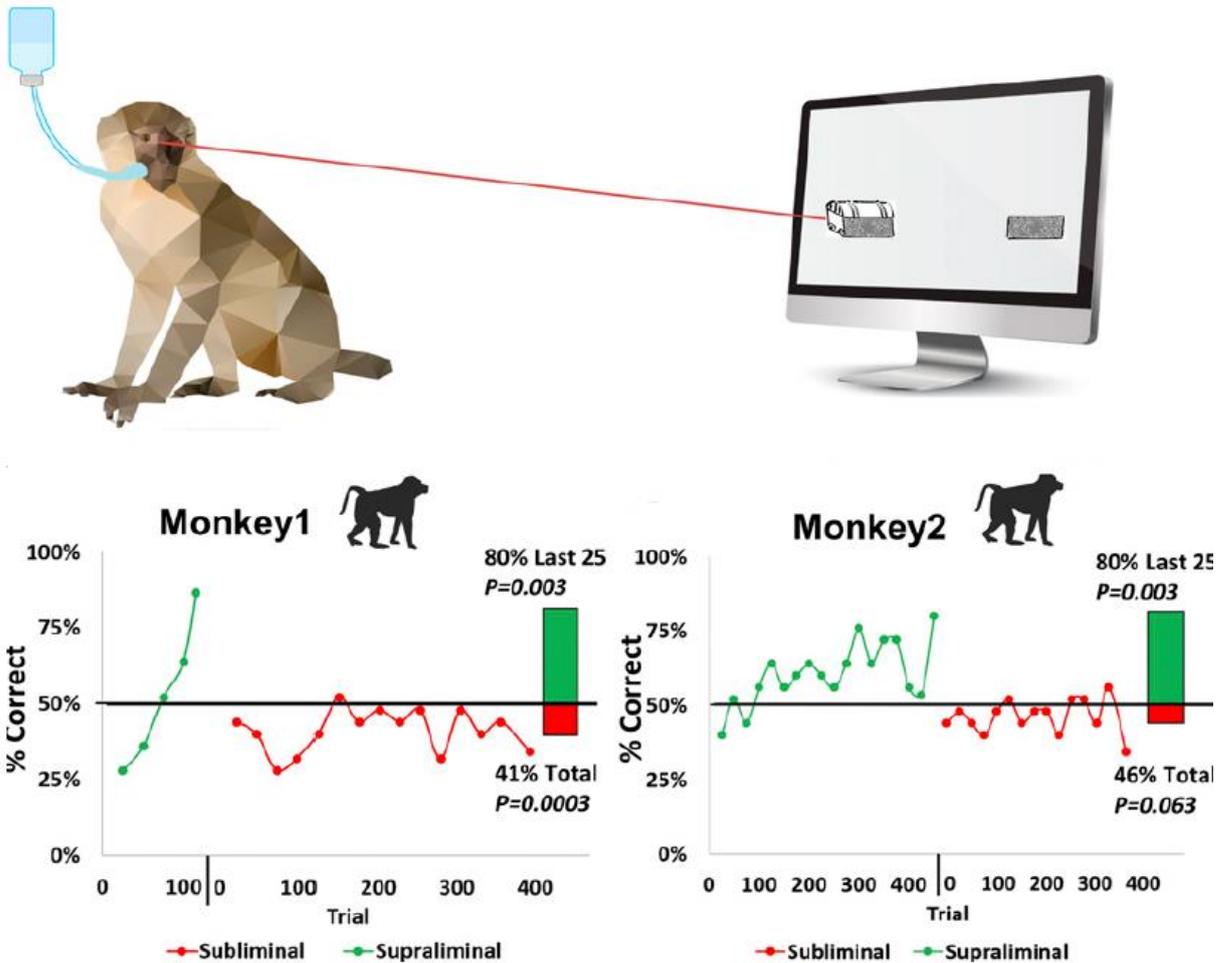


Figura 19 - Ben-Haim et al, 2021, pp. 4-6. Modificado.

Uma possível objeção que os autores levantam seria a de que o desempenho pior dos indivíduos submetidos ao sinal (*cue*) subliminar se deu devido à baixa saliência que tal estímulo possuiria. Em outras palavras, a diferença de desempenho é explicada através do estímulo ser mais ou menos forte e não em virtude de o estímulo ter sido ou não percebido conscientemente. Como Crump e Birch (2021) bem apontam, o problema de confundir a força do sinal com a percepção consciente do estímulo é algo bastante presente na ciência da consciência (LAU, 2011). Portanto, seria crucial respondê-la para garantir a confiabilidade do experimento. Para testar essa hipótese contrária, os pesquisadores realizaram 2 experimentos.

No primeiro experimento, os pesquisadores informaram a alguns participantes que eles seriam apresentados a um sinal (*cue*). Em razão disso, muitos dos participantes relataram ter visto o sinal (*cue*) e, assim, obtiveram um desempenho melhor do que aqueles que relataram não o ter visto. Ou seja, apesar dos sinais (*cues*) não terem aumentado a sua força, apenas o fornecimento da informação de que seria exibido um sinal (*cue*) melhorou o desempenho

daqueles que relataram o ter visto. Isto sugere que é em função da ciência do estímulo, e não a sua força, que resulta na diferença de desempenho dos participantes (BEN-HAIM et al, 2021).

No segundo experimento, os pesquisadores informaram de antemão aos participantes que o sinal (*cue*) predizia a localização do estímulo alvo. Deste modo, os participantes não necessitavam realizar o aprendizado de associação do estímulo incongruente com a localização oposta do estímulo alvo. Neste caso, não seria envolvido o fator aprendizagem. Ainda assim o resultado obtido foi que os indivíduos que relataram ter percebido o sinal (*cue*) obtiveram um desempenho melhor do que aqueles que relataram não ter percebido (BEN-HAIM et al, 2021).

A conclusão retirada desse experimento foi que a percepção consciente do sinal (*cue*) de fato facilitava o desempenho da tarefa de localizar o estímulo alvo. Enquanto a percepção inconsciente tornava mais lento o processo de aprendizagem do sinal (*cue*) com a localização do estímulo alvo (BEN-HAIM et al, 2021; CRUMP e BIRCH, 2021, 2022).

Em resumo, a abordagem teórica leve segue o padrão da abordagem epistêmica de primeiro investigar em humanos e depois nos animais o conjunto (*cluster*) C_m de habilidades cognitivas facilitadas pela consciência. Como dito no início da seção, a ideia aqui não é buscar um único marcador em humanos e verificar se o animal o possui. Tampouco compor uma lista de mercado composta por marcadores verificados em humanos para checar se os animais satisfazem ou não tal lista. Como Birch (2020a, 2021) bem argumenta, devemos buscar um conjunto C_m de marcadores que covariam, pois quanto maior e mais diversificado ele for, mais robusto enquanto conjunto de evidências ele será. Claro, alguns marcadores que compõem C_m poderão estar ausentes nos animais, como, por exemplo, o relato verbal. No entanto, isso não constitui um problema à abordagem teórica leve. Basta que o animal possua, nos termos de Birch (2020a, p. 11), uma “*fração substancial*” de C_m .

3.6.1 Respostas a Possíveis Objeções

No que se segue, buscarei responder algumas das críticas que podem ser levantadas contra a abordagem teórica leve. A primeira crítica que Birch (2020a) suspeita seria a de que talvez todas essas habilidades cognitivas pudessem ocorrer sem a consciência nos animais, embora sejam facilitadas pela consciência nos humanos. Ou como o crítico poderia articular: “[v]ocê mostrou que as habilidades estão *presentes*, porém você não mostrou as suas *facilitações pela consciência*.”¹⁷⁶ (BIRCH, 2020a, p. 11). Birch responde que devemos buscar

¹⁷⁶ “[y]ou’ve shown the abilities are *present*, but you haven’t shown their *facilitation by consciousness*.”

protocolos que gerem percepções *inconscientes* nos animais. Novamente podemos notar o apelo de Birch à metodologia dos contrastes mínimos, amplamente utilizada na ciência da consciência como forma de distinguir processos conscientes dos inconscientes (DEHAENE, 2014; HARLEY, 2021). Dentre os protocolos citados por Birch tem-se tarefas que distraem, *backward masking*, piscar atencional (*attentional blink*) e supressão de flash (*flash suppression*). Assim como Birch, chamarei todos esses protocolos de “mascaramentos”. Grosso modo, as metodologias de mascaramento consistem em tornar uma figura alvo invisível através da sua projeção em um dado tempo mínimo t no meio de uma sequência de formas de imagens similares que agem como “mascadores” da figura alvo (DEHAENE, 2014, pp. 37-41).

Segundo Birch, devemos verificar se nos animais, tal como ocorre nos humanos, a utilização do procedimento de mascaramento “desliga” ou “liga” o C_m (BIRCH, 2020a; CRUMP e BIRCH, 2021, 2022). Citando o exemplo do condicionamento de rastreo, quando o som é exibido de forma subliminar ao ser humano, o condicionamento de rastreo é “desligado”. Por sua vez, quando o mesmo procedimento é feito com o condicionamento com *delay*, ele não é “desligado”, permanecendo intacto (BIRCH, 2020a, pp. 11-12). Ou seja, quando exposto ao estímulo de modo subliminar, o indivíduo realiza o condicionamento com *delay*, mas não o condicionamento de rastreo. Devemos agora verificar se todo o C_m desliga quando submetido ao mascaramento, dado a relação de covariação entre os marcadores que compõem o conjunto (*cluster*). Caso o mesmo ocorra entre os animais, teríamos uma maior robustez em relação à hipótese de que eles são ou não conscientes.

Todavia, o crítico poderia contra-argumentar: “[c]omo você poderia demonstrar que o estímulo foi mascarado com sucesso, dado a ausência do relato verbal?”¹⁷⁷ (BIRCH, 2020a, p. 12). Birch responde da seguinte forma. Imagine que tenhamos descoberto um marcador M_1 e o seu protocolo correspondente de mascaramento que o desliga. O próximo passo consistiria em verificar se o mesmo protocolo também desligaria todos os outros marcadores associados ao conjunto C_m em que ele está inserido. Do mesmo modo, também deveríamos verificar se as habilidades que demandam menos processamento cognitivo, como o condicionamento com *delay*, permanece ou não inalterado. Caso a resposta fosse positiva, isto é, o protocolo de mascaramento somente desligasse ou ligasse os marcadores associados em C_m , deixando inalterados a situação em tela dos processos que demandam menos, isso apoiaria a evidência de que as habilidades de C_m precisam do suporte da consciência para serem executadas. Ademais, isso também garantiria a evidência de que o protocolo de fato mascarou

¹⁷⁷ “[h]ow could you ever show that a stimulus had been successfully masked, in the absence of verbal report?”

o estímulo (BIRCH, 2020a, p. 12). Ao fim e ao cabo, tudo isso forneceria evidências suficientes para justificar a IME acerca da hipótese de que o animal *S* possui ou não estados mentais conscientes (BIRCH, 2020a).

Uma outra objeção que poderia ser feita seria por meio do chamado “Cânone de Morgan” (*Morgan’s Canon*). Creio que esta seja uma objeção que todas as formas de abordagens epistêmicas devem enfrentar. Infelizmente Birch (2020a) não a avalia. No que se segue, responderei tal objeção de modo a assegurar a abordagem teórica leve.

O Cânone de Morgan (CM) é um princípio formulado pelo psicólogo britânico C. L. Morgan, um dos fundadores da psicologia comparada moderna (MORGAN, 1903). Ainda hoje este princípio desempenha uma grande influência no debate acerca do estatuto da mente e do comportamento animal (ANDREWS, 2020a, 2020b; FITZPATRICK, 2018; SHETTLEWORTH, 2010, 2012; WYNNE e UDELL, 2020). O CM afirma que:

*Em caso algum devemos interpretar uma ação como o produto do exercício de uma faculdade psíquica superior, se ela pode ser interpretada como o produto do exercício daquele que se encontra no nível inferior da escala psicológica.*¹⁷⁸ (MORGAN, 1903, p. 53)

Existem divergências na literatura de como devemos interpretar os termos “inferior” e “superior” que o CM assume. Seguindo Simon Fitzpatrick (2018), interpretarei ambos como “não cognitivo” e “cognitivo” respectivamente. Os processos cognitivos seriam aqueles que utilizam estados mentais, como: emoções, crenças, desejos, representações entre outras, enquanto os não cognitivos seriam aqueles compostos por apenas processos associativos, como os que fazem parte do condicionamento clássico e operante (assumindo que tais processos não utilizam estados mentais). O princípio do CM pode ser interpretado de diversas formas (FITZPATRICK, 2018). Aqui me restringirei a responder somente a versão que Fitzpatrick (2018) denominou de “O Cânone Restrito” (a partir deste momento usarei os termos “Cânone Restrito” e “Cânone de Morgan” de forma sinônima, a menos que seja apontado o contrário). Segundo esta versão, devemos interpretar o comportamento de um dado animal por meio de explicações de nível inferior até que as evidências as descartem por completo (FITZPATRICK, 2018, p. 439). Em contraste com uma interpretação mais forte, ela nos diz que não é que devemos aceitar automaticamente explicações de nível inferior quando

¹⁷⁸ “*In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale.*”

os dados comportamentais são compatíveis com ambos os níveis, mas sim: “[q]uando uma explicação de nível inferior se encontra disponível, nós não devemos aceitar a de nível superior.”¹⁷⁹ (FITZPATRICK, 2018, p. 439). Em outras palavras, só devemos aceitar as explicações de nível superior quando as de nível inferior forem completamente descartadas.

Essa versão do CM indicaria que, a não ser que seja completamente rejeitada, devemos sempre escolher a explicação de nível inferior quando ela se encontra disponível. Isto impactaria de maneira direta as abordagens epistêmicas, pois mesmo que exista um conjunto de evidências que sustente uma explicação cognitiva do comportamento do animal em questão, se existe uma explicação não cognitiva do mesmo comportamento, ela deverá ser escolhida em detrimento da outra. Por exemplo, suponha que seja observado um gato preso dentro de uma jaula que ao aprender a abri-la, vai em direção à comida que se localizada no lado de fora. Ou imagine um crocodilo exibindo um comportamento que denota a capacidade de realizar o condicionamento de rastreio. Segundo o CM, se existe uma interpretação não cognitiva destes comportamentos, então deve-se interpretá-los dessa forma, mesmo que haja evidências robustas que sustentem uma explicação cognitiva.

Além dessa interpretação do CM possuir problemas sérios, como aponta Fitzpatrick (2018), creio que ela não refuta a hipótese da facilitação. A abordagem teórica leve emprega um conjunto C_m de marcadores de consciência que covariam e assim garantem a robustez e justificção para IME ser utilizada pela melhor hipótese que indica a presença da consciência nos animais. Caso o animal S apresente C_m , dado esse conjunto de evidências robustas que apresentam uma certa relação de covariação, podemos via IME identificar a presença ou não da consciência em S . Neste caso, a melhor explicação do comportamento C seria a consciência, dado o conjunto de evidências, as virtudes teóricas e a capacidade explanatória que a apoiam. A tentativa de explicação de C apenas por meio de meras associações não satisfaria a descrição do evento, pois o fenômeno seria melhor explicado utilizando a consciência. O emprego de apenas associações deixaria de fora uma série de fatores relevantes sobre o comportamento do animal. Ora, como explicar a forma de aprendizado que ocorre no condicionamento de rastreio sem o emprego da consciência, quando há uma diferença substancial entre os comportamentos exibidos quando os estímulos são apresentados supraliminarmente e subliminarmente? Ademais, não faria sentido explicarmos o mesmo comportamento em humanos recorrendo à consciência como o seu facilitador, mas sob condições idênticas ou semelhantes e explicarmos o comportamento do animal sem o uso da consciência (mesmo que o animal apresente C_m). Por

¹⁷⁹ “[w]hen a lower explanation is available, we must not accept a higher one.”

qual razão a seleção natural faria com que *C* fosse causado pela consciência nos humanos, mas em outros animais não?¹⁸⁰ A explicação de *C* através do uso exclusivo de associações deixaria lacunas que somente uma hipótese robusta e com evidências suficientes preencheria.

No entanto, poderia ser objetado que a história evolutiva dos outros organismos seria tão diferente da nossa que o surgimento da consciência entre eles pode não ter ocorrido através do seu favorecimento pela seleção natural. Ou seja, seria plausível considerar a hipótese de que apesar da consciência ter sido favorecida pela seleção natural nos humanos, o mesmo não ocorreu nos outros animais. Assim, por terem histórias evolutivas bastante diferentes das nossas, a consciência não surgiu entre os animais. Note que essa objeção pode ser orientada por duas visões distintas: uma adaptacionista e outra “gouldineana”¹⁸¹. A primeira visão sustentaria que a consciência não surgiu nos animais porque diferentemente dos humanos ela não foi favorecida pela seleção natural. A segunda visão, por sua vez, expressaria que em virtude de a consciência ser um subproduto do traço ou conjunto de traços *T* nos humanos e pelos animais não possuírem *T*, a consciência não surgiu entre os animais. Em outras palavras, pela consciência ser um *spandrel* oriundo de um traço ou conjunto de traços *T* nos humanos, e nenhum outro animal possuir *T*, então os animais não poderiam ser dotados de consciência. A consciência, deste modo, seria uma característica exclusivamente humana. Para responder de forma adequada esta objeção, a dividirei em duas questões distintas. A primeira delas tocaria os animais filogeneticamente próximos dos humanos, enquanto a segunda os mais distantes. As responderei concisamente a seguir.

No que tange aos animais filogeneticamente próximos de nós, como os outros homínídeos, ou mesmo todos os mamíferos, não considero razoável que a história evolutiva deles seja tão diferente a ponto de a seleção natural não ter favorecido o surgimento da consciência. Compartilhamos com os outros mamíferos e as aves muitos dos comportamentos, da neurofisiologia, da cognição, dos estados perceptuais e afetivos (CHURCHLAND, 2013, pp. 248-249). Não me soaria estranho se a consciência fosse um traço homólogo entre todos, ou quase todos, os mamíferos. Isto é, a consciência seria um traço compartilhado entre todas as espécies de mamíferos, em razão de ser produto de um ancestral em comum que, por sua vez, também a possuía (FUTUYMA, 2009; RIDLEY, 2013). Ademais, a maior parte dos mamíferos ser dotada de consciência é, com raras exceções, um ponto pacífico na literatura (BAARS,

¹⁸⁰ Com isso não quero me comprometer com nenhuma forma de adaptacionismo, tampouco com a hipótese de que a consciência seja um *spandrel*.

¹⁸¹ Em referência ao paleontólogo Stephen J. Gould, um dos autores do clássico artigo contra o adaptacionismo (GOULD e LEWONTIN, 1979).

2005a; CHURCHLAND, 2013; DEHAENE, 2014; FEINBERG e MALLATT, 2017; GINSBURG e JABLONKA, 2019; GODFREY-SMITH, 2020a; LAU, 2022; TOMASELLO, 2022).¹⁸² Deste modo, seria evolutivamente mais parcimonioso considerar a consciência um traço compartilhado entre os mamíferos, ou ao menos entre os hominídeos, do que considerá-la um milagre que surgiu apenas recentemente entre os humanos (SOBER, 2000). Em outras palavras, é evolutivamente mais parcimonioso e explanatoriamente simples considerar que o mesmo comportamento *C* foi causado pela consciência em humanos e entre os mamíferos do que o contrário. Ora, se ambos os humanos e os mamíferos engajam em *C* e possuem estruturas neurofisiológicas e comportamentos semelhantes, e exibem um conjunto de marcadores de consciência que covariam, soa implausível e *ad hoc* eliminar a consciência como a explicação da causa de *C* nesses animais ao mesmo tempo que se admite nos humanos.

Ainda que a consciência seja um fenômeno exclusivo entre os humanos, não me parece claro sobre exatamente de *quais* humanos estaríamos tratando. Entrariam neste grupo os *Homo erectus*, espécie pertencente ao nosso gênero onde a cerca de 1.5 milhão de anos atrás populações do sudeste asiático e da África possuíam uma organização cerebral, incluindo a do lobo frontal, similar ao do *Homo* moderno (NETO et al, 2022; de LEÓN et al, 2021)? Ou mesmo os *Australopithecus*, que estão entre os hominínios mais arcaicos descobertos e que provavelmente já fabricavam ferramentas rudimentares, como as que pertencem à indústria Olduvaiense (GRATÃO et al, 2022; NETO et al, 2022; PARINS-FUKUSHI et al, 2019)? Existem evidências sólidas de que os *Homo neanderthalensis* produziram até mesmo pinturas rupestres, o que poderia denotar a presença mesmo que primitiva de pensamento simbólico (HOFFMANN et al, 2018; MARTÍ et al, 2021). Embora seja polêmico, alguns autores através dos registros arqueológicos relativos à morfologia, como a presença de estruturas vocais quase modernas e à cultura, como a construção de ferramentas, edifícios e a realização de sepultamentos ritualizados, inferem que os neandertais possuiriam a capacidade de linguagem (mesmo que primitiva) (DEDIU e LEVINSON, 2018).¹⁸³ Seriam os neandertais também conscientes? Ou por humanos os autores céticos gostariam de se referir apenas ao *Homo sapiens*

¹⁸² Em uma pesquisa recente realizada com pesquisadores durante os encontros anuais de 2018 e 2019 da Associação do Estudo Científico da Consciência, a maior parte dos entrevistados consideraram que mamíferos, polvos e até peixes são provavelmente seres conscientes, ao passo que minhocas, árvores, pedras e termostatos não os são (FRANCKEN et al, 2022, p. 8).

¹⁸³ Outros autores, como Planer e Sterelny (2021) e Botha (2020), são mais cautelosos. Planer e Sterelny (2021, pp. 178-179), por exemplo, sustentam que apesar dos neandertais satisfazerem alguns dos pré-requisitos para que ocorra linguagem, como as estruturas vocais, auditivas e cognitivas, eles possuiriam apenas uma forma de “protolinguagem”. O motivo, segundo eles, é que a linguagem moderna depende não só de estruturas individuais, mas também de características relacionadas ao ambiente social, algo que os neandertais careceriam. Berwick e Chomsky (2017, p. 177), por sua vez, são céticos quanto à presença de linguagem em neandertais.

que surgiu entre 300 e 200 mil anos atrás (GALWAY-WITHAM e STRINGER, 2018)? Porém, mesmo entre os *Homo sapiens* os comportamentos que supostamente denotariam pensamento simbólico, como a produção de ornamentos, artes rupestres e sepultamentos ritualizados só surgiram há pelo menos 50 mil anos atrás no Paleolítico Superior durante o chamado “Grande Salto Adiante” (DIAMOND, 2020, 2021).¹⁸⁴ Seriam os humanos anteriores ao Grande Salto Adiante destituídos de consciência? Quanto a essas perguntas, os céticos se mantêm em silêncio (CARRUTHERS, 2019).

Portanto, além de não ser parcimonioso, o argumento do milagre do surgimento da consciência soaria como uma forma de antropocentrismo descabido. Digo antropocentrismo no sentido de que uma vez que a consciência é uma característica especial, ela só poderia ter surgido em uma espécie tão especial quanto. Como a espécie humana é especial, seja pelo motivo que for (e. g. linguagem, pensamento altamente abstrato, organização social complexa), logo a consciência só poderia ter surgido entre os humanos. Para além dos já viesados, não creio que esse argumento convença alguém atento as evidências atuais.

No entanto, o apelo à parcimônia evolutiva se torna frágil conforme se chega em clados mais distantes do nosso. Como então responder a objeção do surgimento da consciência quando ela se refere à grandes distâncias evolutivas, como a que separa os humanos e os cefalópodes, cujo último ancestral em comum é de cerca de 550 milhões de anos atrás (Figura 20) (GODFREY-SMITH, 2019a; SCHNELL et al, 2020)? Acredito que a melhor forma de responder a essa objeção seria através do processo conhecido como “evolução convergente”. Neste processo, poderia ser o caso de dois ou mais traços que guardam semelhanças importantes surgirem independentemente em duas ou mais espécies. Em outras palavras, é plausível que surjam traços que compartilham similaridades relevantes, como certas especificidades anatômicas, mas que evoluíram de forma independente entre duas ou mais espécies, não sendo, portanto, provenientes de um mesmo ancestral em comum. A esse traço dá-se o nome de “homoplasia” (FUTUYMA, 2009; RIDLEY, 2013). Um exemplo paradigmático de homoplasia

¹⁸⁴ Algumas evidências apontam que tal evento poderia ter ocorrido entre 250 e 80 mil anos atrás, porém em menor escala do que ocorreu a cerca de 50 mil anos atrás (BOUZOUGGAR et al, 2007; MCBREARTY e STRINGER, 2007). Ou seja, provavelmente a maior abundância de material simbólico encontrada no registro fóssil de 50 mil anos atrás foi devido a intensificação e propagação das práticas simbólicas, e não a uma mudança abrupta na cognição dos hominínios durante esse período (STERELNY, 2021, pp. 22-23). As razões que teriam levado ao surgimento do comportamento moderno dos humanos é também matéria de um longo e controverso debate. Alguns autores apontam que o surgimento do comportamento moderno foi devido à complexificação da cognição, via, por exemplo, a memória de trabalho (WYNN e COOLIDGE, 2011). Outros, por sua vez, consideram que foi em razão da capacidade de transmissão confiável de informações ao longo das gerações (STERELNY, 2011). Independentemente se as práticas que denotam pensamento simbólico surgiram a cerca de 250 ou 50 mil anos atrás, o presente argumento se mantém intacto, pois ainda estaríamos tratando de hominínios cuja arquitetura cognitiva provavelmente se difere bastante dos humanos neurotípicos atuais.

é a nadadeira dorsal dos golfinhos e dos peixes. Apesar de serem anatomicamente e funcionalmente parecidas, elas não são produtos de um ancestral em comum. Essas características evoluíram de forma independente nas duas espécies. Retomando o caso anteriormente citado dos humanos e cefalópodes, considere o seguinte breve exemplo.

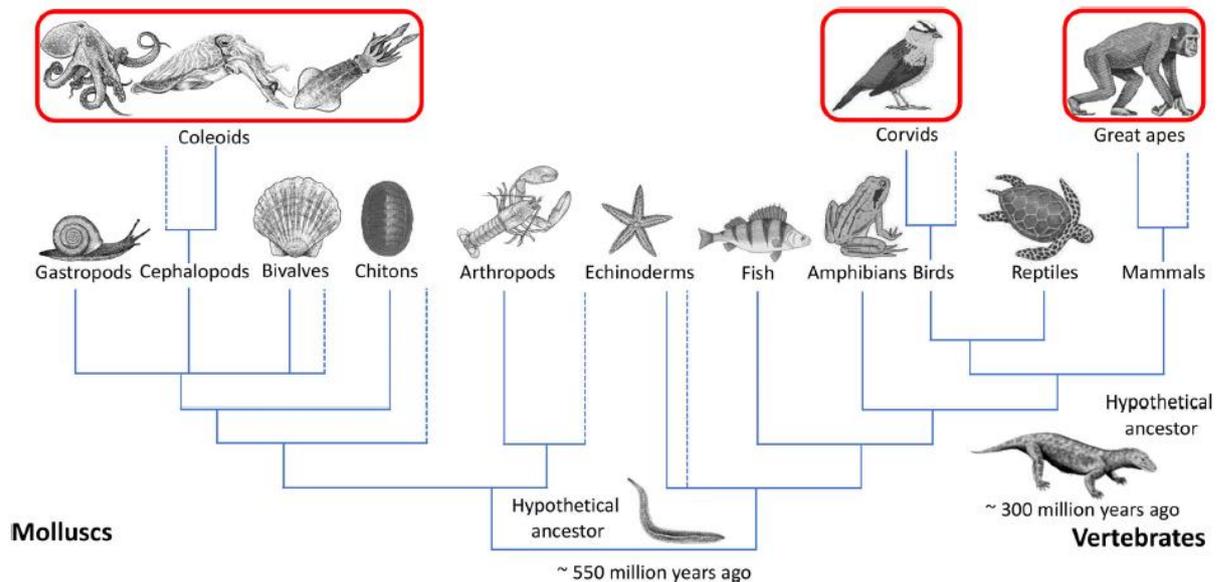


Figura 20– A árvore filogenética sobre as relações evolutivas entre diversos clados, especialmente os coleóides (subgrupo dos cefalópodes), corvídeos (aves) e os grandes primatas (mamíferos). Schnell et al, 2020, p. 165.

No que toca a evolução dos olhos, Dan-E. Nilsson (2013) argumenta que existiriam ao menos quatro classes de olhos, sendo a classe I a mais primitiva e simples e a IV a mais recente e complexa (Figura 21). Segundo o autor, enquanto os olhos da classe IV proporcionam uma alta resolução visual, os de classe I proporcionam uma baixa resolução aos organismos. Os olhos da classe IV podem ser distinguidos entre duas formas: o tipo câmara e o tipo composto (*compound*). São poucos os animais que desenvolveram esta classe complexa de olhos. Dentre eles pode-se citar os vertebrados, os artrópodes e os cefalópodes. O interessante aqui a se notar é o seguinte: os olhos de classe IV do tipo câmara evoluíram independentemente entre os humanos e os cefalópodes. Ou seja, embora os olhos desses grupos possuam uma história evolutiva diferente, ambos chegaram em formas modernas que guardam semelhanças relevantes. Claro, os nossos olhos e os dos polvos, por exemplo, possuem algumas características bastante diferentes. Os polvos (*Octopus vulgaris*), por exemplo, são *color-blind*, isto é, eles são incapazes de verem cores em razão de possuírem somente um único pigmento visual em seus fotorreceptores (HANKE e KELBER, 2020). Isso é no mínimo surpreendente, pois esses organismos são capazes de gerar camuflagens cujas cores são extremamente precisas em relação às cores, texturas e outros detalhes presentes no ambiente que tentam simular. De

modo geral, no entanto, ambos os olhos dos humanos e dos polvos compartilham características bastante similares, como as de natureza fisiológica e funcional. Tanto os olhos dos humanos quanto os dos cefalópodes possuem, por exemplo, retina e cristalino. Ora, possivelmente estaríamos diante de uma homoplasia. Neste caso, os olhos de classe IV do tipo câmera evoluíram independentemente nos humanos e nos polvos, uma vez que não surgiram de um ancestral em comum.¹⁸⁵ Isso é bastante notável, pois não só possuímos histórias evolutivas completamente distintas, como o nosso último ancestral em comum se encontra a uma distância temporal significativamente longínqua (por volta de 550 milhões de anos).

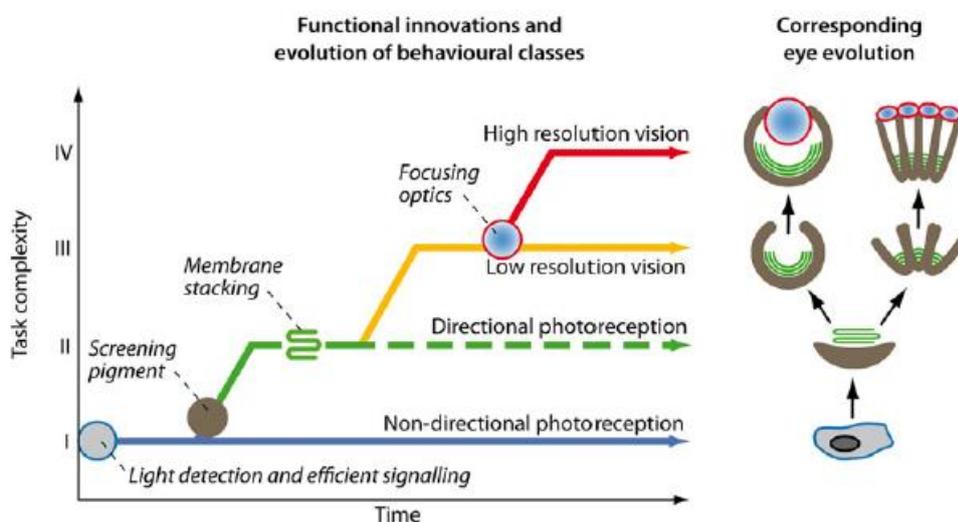


Figura 21 – Evolução dos olhos. Nilsson, 2013, p.10.

Isto posto, considero plausível que o mesmo processo possa ter ocorrido com a consciência. Creio que por princípio nada impede a consciência de ter surgido independentemente e mais de uma vez em linhagens distintas, cujo último ancestral em comum fosse temporalmente distante e possuísse características morfológicas bastante diferentes dos organismos pertencentes às linhagens modernas. À luz do conjunto de evidências de natureza distintas disponíveis, como as neurofisiológicas e as comportamentais, alguns pesquisadores já admitem a possibilidade de a consciência ter surgido através da evolução convergente, como, por exemplo, em uma série de linhagens amniotas (NIEDER, 2021). Ora, dado essa coleção de

¹⁸⁵ Poder-se-ia dizer mais especificamente que esse é um caso de homoplasia do tipo evolução convergente, em contraposição com a evolução em paralelo (FUTUYMA, 2009, pp. 110-111). No entanto, não só há uma profusão de definições diferentes sobre esses conceitos, como existe um debate sobre se de fato faz sentido tal distinção. Alguns autores, por exemplo, argumentam que a distinção entre evolução convergente e paralela é uma falsa dicotomia (ARENDRT e REZNICK, 2008). Para o presente argumento não será necessário admitir ou negar tal distinção.

evidências, pode-se considerar razoável a hipótese de que algo similar ocorreu com certos grupos de invertebrados (como os próprios coleóides supracitados). A consciência, portanto, pode ter sido a característica que mais obteve sucesso na resolução de certos problemas impostos pelo ambiente, contribuindo de forma relevante ao *fitness* do organismo que a possuía.

Alguns autores argumentam que, por exemplo, a consciência teria surgido a cerca de 540 milhões de anos atrás durante o período Cambriano como um traço que desempenhou um papel fundamental na relação de predação. Isso teria ocorrido inicialmente nos predadores através da complexificação de seus sistemas nervosos e de suas cognições (FEINBERG e MALLATT, 2017, 2018; GODFREY-SMITH, 2019b).¹⁸⁶ Seria plausível considerar que a consciência tenha surgido como um *facilitador* da execução de comportamentos e processos cognitivos complexos. Ora, assim como os olhos de classe IV tipo câmera são os traços que evolutivamente mais geraram sucesso às certas necessidades de determinados organismos, a consciência seria o traço que gerou mais sucesso a certas necessidades dos organismos que a possuíam. Claro, essas são hipóteses que merecem uma investigação científico-filosófica mais adequada. Contudo, acredito que apresentam uma boa ilustração do argumento aqui fornecido.

Resumindo o que aqui foi proposto, no que diz respeito à objeção do porquê a seleção natural favoreceu que o comportamento C fosse causado pela consciência nos humanos, mas não em animais com histórias evolutivas distintas, busquei responder de duas maneiras. A primeira foi a seguinte. No que se refere os animais filogeneticamente próximos de nós, seria evolutivamente mais parcimonioso considerar a hipótese de que esses animais também possuem consciência e que os seus comportamentos C foram causados por ela, do que a hipótese contrária. Enquanto em relação aos animais filogeneticamente mais distantes de nós, avaliei que a hipótese de a consciência ser uma homoplasia é a melhor resposta. Neste caso, a consciência pode ter surgido independentemente entre linhagens distintas em razão de algum motivo x , como, por exemplo, ser a característica mais vantajosa para o sucesso reprodutivo e de sobrevivência do organismo y frente às pressões daquele dado ambiente. Ou, de forma mais geral: a seleção natural favoreceu o aumento da frequência do traço T da consciência em certas populações de organismos O , pois foi T o que forneceu mais sucesso a O frente às pressões seletivas P_n impostas por um ambiente A .

Retomando a objeção do Cânone Restrito, ao fim e ao cabo o considero ser uma espécie de *princípio da pior explicação*. Suponha que o conjunto de evidências C_e justifique e corrobore a hipótese H_c de que o animal S possua consciência. Por qual motivo deveríamos

¹⁸⁶ Para um resumo crítico da posição de Feinberg e Mallatt (2018), cf. Barcellos (2021).

seguir o C none Restrito e escolher uma explica o que utiliza apenas associa es (chamo de Q) e que n o satisfaz a explica o demandada pela complexidade do fen meno? Parece-me que a escolha de H_q , a hip tese que utiliza apenas Q , ao inv s de H_c seria uma prefer ncia acr tica pela primeira explica o. Diante deste caso, a melhor hip tese seria a que utiliza H_c e n o H_q como a sua explica o. A afirma o do C none Restrito   nada mais que um dogma, uma vez que ela expressa que mesmo que haja evid ncias robustas corroborando H_c , a escolha deve ser sempre a favor da hip tese que cont m apenas associa es em seu conjunto de evid ncias. Em outras palavras, mesmo que H_c consista na melhor explica o de C , devemos sempre escolher H_q , mesmo que ela resulte em uma explica o insuficiente do fen meno. Uma vez convertido em dogma, o C none Restrito perverteria a no o de escolha da hip tese mais simples para explicar um determinado fen meno. Evidentemente, n o   isso que se deve buscar em uma investiga o criteriosa, rigorosa e aberta a forma diferentes de evid ncias. Deste modo, considero razo vel que se rejeite o uso do C none Restrito nas investiga es relativas   consci ncia em animais. Como foi visto, existe dispon vel uma gama de concorrentes ao CM que s o mais sofisticados e promissores para o uso em pesquisa. Este   o caso da abordagem te rica leve que penso ser um  timo ponto de partida para as investiga es cient fico-filos ficas sobre a consci ncia animal.

CONCLUSÃO

Ao longo desta dissertação, busquei lidar com dois problemas centrais da filosofia e da ciência da consciência animal, a saber, o problema da distribuição e o problema da mensuração. Como foi visto, o problema da distribuição se pergunta até onde a consciência se estende na árvore filogenética (se é que ela realmente se estende para além dos humanos). O problema da mensuração, por sua vez, trata de qual seria o método correto para atribuir consciência aos animais. Os problemas se conectam uma vez que, para determinar até onde a consciência se estende na árvore filogenética, devemos ter um método capaz de indicar a sua presença. São três as abordagens que buscam responder esses problemas. A primeira delas é a abordagem da percepção direta. Grosso modo, ela sustenta que de alguma forma percebemos os estados conscientes dos animais de forma direta, isto é, sem a mediação de inferências em nossa mente. Já a abordagem teórica sustenta que é a aplicação de uma teoria da consciência humana no animal em questão que determina se ele possui estados mentais conscientes. Por fim, a abordagem epistêmica sustenta que é através da estratégia dos marcadores de consciência que se justifica a hipótese sobre se um animal é ou não consciente.

No início do 1º capítulo, busquei elucidar o uso que faria do termo “consciência”. Como Block (1995) o fez, pontuei que “consciência” é um conceito ambíguo e difícil, se não impossível, de se definir. Por acreditar serem as formas metafisicamente menos onerosas, empreguei, assim como Schwitzgebel (2016), a estratégia de utilizar exemplos e a locução “há algo para alguém que é como ser *x*” de Nagel (1974) como forma de explicitar o conceito de “consciência”. Posteriormente, realizei algumas distinções importantes sobre os diferentes sentidos que “consciência” carrega na literatura. Também forneci um breve panorama do debate sobre a possível distinção empírica entre a consciência fenomenal e a consciência de acesso, particularmente sob o prisma do argumento do transbordamento (BLOCK, 2007a).

No 2º capítulo, apresentei criticamente a posição cético-radical de Peter Carruthers (2019). De acordo com Carruthers, não haveria questão de fato sobre a atribuição de consciência aos animais. Os juízos que atribuem consciência aos animais seriam semanticamente indeterminados. A ciência da consciência animal consistiria, portanto, em um projeto fadado ao fracasso. A sugestão de Carruthers (2019, p. 164) é a de que os esforços científicos-filosóficos deveriam se concentrar apenas na investigação dos processos cognitivos dos animais. Forneci argumentos contrários à posição de Carruthers que considero serem suficientes para justificar o seu abandono. Ao contrário do que Carruthers prega, conclui que

as pesquisas científicas e filosóficas sobre a consciência animal devem continuar a serem realizadas.

No 3º e último capítulo, me aprofundei na abordagem epistêmica. Analisei duas importantes versões dela: a abordagem padrão e a abordagem teórica neutra. Em relação à segunda, debrucei-me especificamente na versão newtoniana de Tye (2017, 2018). Ao final do capítulo, defendi a abordagem teórica leve, posição que foi originalmente proposta por Birch (2020a). Busquei mostrar que a abordagem teórica leve é, dentre as posições recentes, a mais empiricamente adequada e conceitualmente plausível. Também respondi a algumas das possíveis objeções que poderiam ser feitas a ela.

Considerando a vasta quantidade de literatura e do rápido avanço da área, admito ter deixado muitas lacunas neste trabalho. Desejo conseguir preencher ao menos algumas delas em pesquisas futuras. Por fim, espero que a presente dissertação contribua de alguma forma às pesquisas filosóficas e científicas sobre a consciência em animais. Há muito trabalho para ser realizado.

APÊNDICE

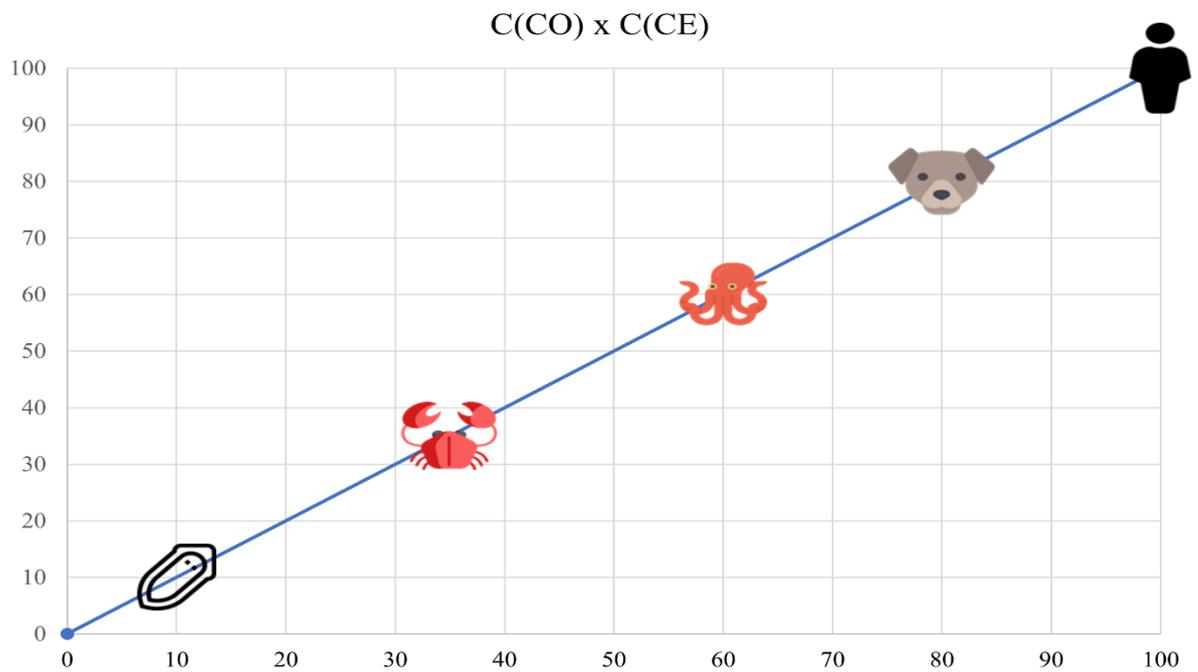


Figura 22 – Representação gráfica do modelo gradualista *scala naturae* da consciência. O eixo x se refere ao nível de complexidade (C) da consciência (CO). O eixo y se refere ao nível de complexidade (C) cerebral (CE). Cada ponto da linha, com exceção do 0, denota um grupo animal diferente. Sendo: pt (10): nematoide; pt (35): crustáceo; pt (60): polvo; pt (80): canídeo; pt (100): humano.

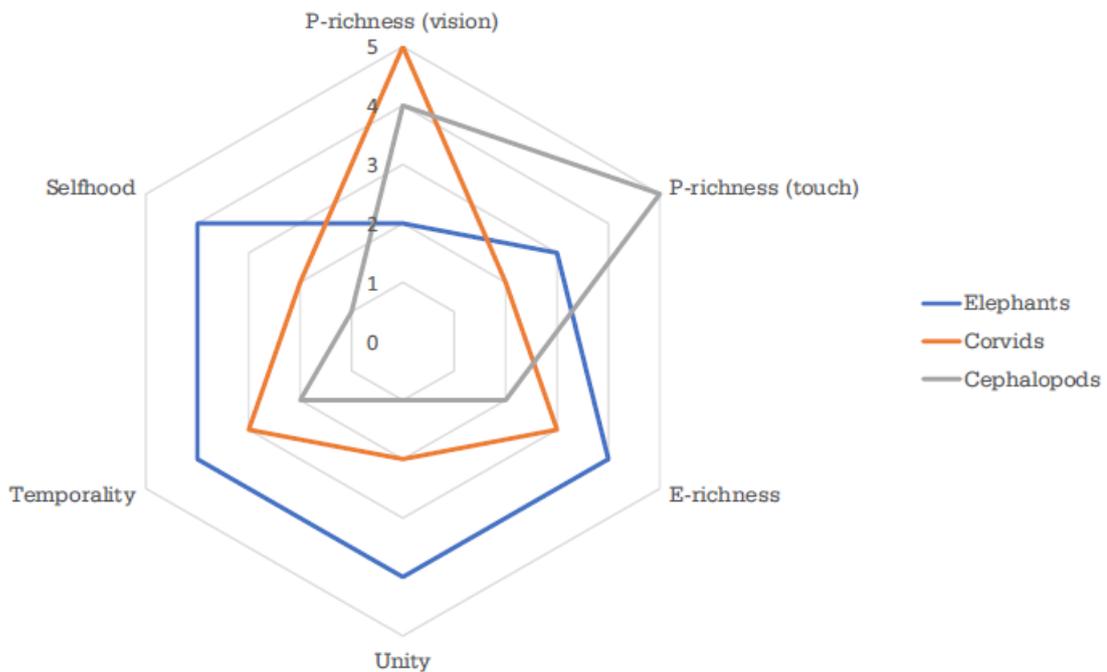


Figura 23 – Representação gráfica do modelo gradualista multidimensional da consciência. Linha azul: Elefantes; Laranja: Corvídeos; Cinza: Cefalópodes. Tradução das dimensões no sentido anti-horário: Riqueza-P (visão) (“P” se refere à “percepção”), Individualidade, Temporalidade, Unidade, Riqueza-A (“E” se refere à “avaliatividade” ou “*evaluative*” no original), Riqueza-P (tato). Birch et al, 2020, p. 791. Adaptado.

REFERÊNCIAS

- ABEN, B. et al. About the Distinction Between Working Memory and Short-Term Memory. *Frontiers in Psychology*, v. 3, pp. 1-9, 2012.
- ALBURQUERQUE, N. et al. Dogs Can Infer Implicit Information from Human Emotional Expressions. *Animal Cognition*, v. 25, pp. 231-240, 2022.
- ALCOCK, J.; RUBENSTEIN, D. R. *Animal Behavior*. 11 ed. New York: Oxford University Press, 2019.
- ALLEN, C. et al. The Lower Bounds of Cognition: What do Spinal Cords Reveal? In: BICKLE, J. (Org.) *The Oxford Handbook of Philosophy and Neuroscience*. New York: Oxford University Press, 2009.
- ALLEN, C. Animal Pain. *Noûs*, v. 38, n. 4, pp. 617-643, 2004.
- _____. Associative Learning. In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018.
- ALLEN, C.; TRESTMAN, M. *Animal Consciousness*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2016 [Acessado em 20 de março de 2021] <https://plato.stanford.edu/archives/win2020/entries/consciousness-animal>
- ANDREWS, K. Mind. In: GRUEN, L. (Org.) *Critical Terms for Animal Studies*. Chicago: The University of Chicago Press, 2018.
- _____. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. 2 ed. New York: Routledge, 2020a.
- _____. *How to Study Animal Minds*. Cambridge: Cambridge University Press, 2020b.
- ANDREWS, K.; MONSÓ, S. *Animal Cognition*. ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2021 [Acessado em 20 de março de 2021]: <https://plato.stanford.edu/archives/spr2021/entries/cognition-animal>
- ARENDET, J.; REZNICK, D. Convergence and Parallelism Reconsidered: What Have We Learned About the Genetics of Adaptation? *Trends in Ecology & Evolution*, v. 23, n. 1, pp. 26-32, 2008.
- ARMSTRONG, D. *A Materialist Theory of the Mind*. 2 ed. New York: Routledge, 1993.
- BAARS, B. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press, 1988.
- _____. Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience. *Progress in Brain Research*, v. 150, pp. 45-53, 2005a.

- _____. Subjective Experience is Probably Not Limited to Humans: The Evidence from Neurobiology and Behavior. *Consciousness and Cognition*, v. 14, n. 1, pp. 7-21, 2005b.
- BAARS, B. et al. Global Workspace Theory (GWT) and Prefrontal Cortex: Recent Developments. *Frontiers in Psychology*, v. 12, pp. 1-6, 2021.
- BACKER, K. C. Introduction to Experimental Methods in Cognitive Neuroscience. In: YOUNG, B. D.; JENNINGS, C. D. (Orgs.) *Mind, Cognition, and Neuroscience: A Philosophical Introduction*. New York: Routledge, 2022.
- BADDELEY, A. D.; HITCH, G. Working Memory. *Psychology of Learning and Motivation*, v. 8, pp. 47-89, 1974.
- BALAKHONOV, D.; ROSE, J. Crows Rival Monkeys in Cognitive Capacity. *Scientific Reports*, v. 7, pp. 1-8, 2017.
- BALL, D. There Are No Phenomenal Concepts. *Mind*, v. 118, n. 472, pp. 935-962, 2009.
- BALOG, K. In Defense of the Phenomenal Concept Strategy. *Philosophy and Phenomenological Research*, v. 84, n. 1, pp. 1-23, 2012.
- BARCELLOS, V. M. O Neuronaturalismo Biológico e a Consciência sob uma Perspectiva Neurobiológica e Evolutiva: Resenha do livro *Consciousness Desmystified*. *Revista Reflexões*, n. 18, pp. 247-254, 2021.
- BARRON, A. B; KLEIN, C. What Insects Can Tell Us About the Origins of Consciousness. *Proceedings of the National Academy of Science United States of America*, v. 113, n. 18, pp. 4900-4908, 2016.
- BARTELS, A. Consciousness: What is the Role of Prefrontal Cortex? *Current Biology*, v. 31, n. 13, pp. 853-856, 2021.
- BATESON, M. et al. Agitated Honeybees Exhibit Pessimistic Cognitive Biases. *Current Biology*, v. 21, n. 12, pp. 1070-1073, 2011.
- von BAYERN, A. M. P. et al. Compound Tool Construction by New Caledonian Crows. *Scientific Reports*, v. 8, pp. 1-8, 2018.
- BAYNE, T. *Philosophy of Mind: An Introduction*. New York: Routledge, 2021.
- BAYNE, T. et al. Are There Levels of Consciousness? *Trends in Cognitive Sciences*, v. 20, n. 6, pp. 405-413, 2016.
- BAYNE, T.; MONTAGUE, M. (Orgs.) *Cognitive Phenomenology*. Oxford: Oxford University Press, 2011.
- BAYNE, T.; SHEA, N. Consciousness, Concepts and Natural Kinds. *Philosophical Topics*, v. 48, n. 1, pp. 65-83, 2020.

- BELLETT, J. et al. Decoding Rapidly Presented Visual Stimuli from Prefrontal Ensembles Without Report Nor Post-Perceptual Processing. *Neuroscience of Consciousness*, n. 1, pp. 1-12, 2022.
- BEN-HAIM, M. S. et al. Disentangling Perceptual Awareness from Nonconscious Processing in Rhesus Monkeys (*Macaca mulatta*). *Proceedings of the National Academy of Sciences of the United States of America*, v. 118, n. 5, pp. 1-9, 2021.
- BERENT, I. *The Blind Storyteller: How We Reason About Human Nature*. New York: Oxford University Press, 2020.
- _____. Can We Get Human Nature Right? *Proceedings of the National Academy of Sciences of the United States of America*, v. 118, n. 39, pp. 1-6, 2021.
- BERMÚDEZ, J. L. *Cognitive Science: An Introduction to the Science of the Mind*. 3 ed. New York: Cambridge University Press, 2020.
- BERWICK, R. C.; CHOMSKY, N. *Por que Apenas Nós? Linguagem e Evolução*. São Paulo: Editora Unesp, 2017.
- BIRCH, J. et al. Dimensions of Animal Consciousness. *Trends in Cognitive Sciences*, v. 24, n. 10, pp. 789-801, 2020.
- BIRCH, J. et al. *Review of the Evidence of Sentience in Cephalopod Molluscs and Decapod Crustaceans*. London, 2021. (London School of Economics and Political Science, relatório)
- BIRCH, J. et al. How Should We Study Animal Consciousness? *Journal of Consciousness Studies*, v. 29, n. 3-4, pp. 8-28, 2022.
- BIRCH, J. The Search of Invertebrate Consciousness. *Noûs*, pp. 1-21, 2020a.
- _____. Global Workspace Theory and Animal Consciousness. *Philosophical Topics*, v. 48, n. 1, pp. 21-37, 2020b. Versão consultada: <http://philsci-archive.pitt.edu/18031/1/GWTpreprint%2819-08-20%29.pdf>
- _____. The Hatching of Consciousness. *History and Philosophy of the Life Sciences*, v. 43, n. 121, pp. 1-12, 2021.
- _____. Materialism and the Moral Status of Animals. *The Philosophical Quarterly*, v. 72, n. 4, pp. 795-815, 2022a.
- _____. Sentience and the Science-Policy Interface. *Animal Sentience*, v. 31, pp. 1-3, 2022b.
- _____. Should Animal Welfare Be Defined in Terms of Consciousness? *Philosophy of Science*, no prelo.
- BLOCK, N. On A Confusion About a Function of Consciousness. *Behavioral and Brain Sciences*, v. 18, n. 2, pp. 227–247, 1995.

- _____. Mental Paint and Mental Latex. *Philosophical Issues*, v. 7, pp. 19-49, 1996.
- _____. Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience. *Behavioral and Brain Sciences*, v. 30, n. 5-6, pp. 481-499, 2007a.
- _____. Overflow, Access, and Attention. *Behavioral and Brain Sciences*, v. 30, n. 5-6, pp. 530-548, 2007b.
- _____. Consciousness and Cognitive Access. *Proceedings of the Aristotelian Society*, v. 108, n. 1, pp. 289-317, 2008.
- _____. Comparing the Major Theories of Consciousness. In: GAZZANIGA, M. S. (Org.) *The Cognitive Neurosciences*. Cambridge, MA: MIT Press, 2009.
- _____. Perceptual Consciousness Overflows Cognitive Access. *Trends in Cognitive Sciences*, v. 15, n. 12, pp. 567-575, 2011a.
- _____. The Higher Order Approach to Consciousness is Defunct. *Analysis*, v. 71, n. 3, pp. 419-431, 2011b.
- _____. Rich Conscious Perception Outside Focal Attention. *Trends in Cognitive Sciences*, v. 18, n. 9, pp. 445-447, 2014.
- _____. Empirical Science Meets Higher-Order Views of Consciousness: Reply to Hakwan Lau and Richard Brown. In: PAUTZ, A.; STOLJAR, D. (Orgs.) *Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness*. Cambridge, MA: MIT Press, 2019.
- BLOOM, P. *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books, 2004.
- BOLY, M. et al. Consciousness in Humans and Non-humans Animals: Recent Advances and Future Directions. *Frontiers in Psychology*, v. 4, pp. 1-20, 2013.
- BOLY, M. et al. Are the Neural Correlates of Consciousness in the Frontal or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of Neuroscience*, v. 37, n. 40, pp. 9603-9613, 2017.
- BORAN, E. et al. Persistent Neuronal Firing in the Medial Temporal Lobe Supports Performance and Workload of Visual Working Memory in Humans. *NeuroImage* v. 254, pp. 1-11, 2022.
- BOSTROM, N. Are You Living in a Computer Simulation? *Philosophical Quarterly*, v. 53, n. 211, pp. 243-255, 2003.
- BOTHA, R. *Neanderthal Language: Demystifying the Linguistic Powers of our Extinct Cousins*. Cambridge: Cambridge University Press, 2020.

- BOUZOUGGAR, A. 82,000-year-old Shell Beads from North Africa and Implications for the Origins of Modern Human Behavior. *Proceedings of the National Academy of Sciences of the United States of America*, v. 104, n. 24, pp. 9964-9969, 2007.
- BRONFMAN, Z. Z. et al. We See More Than We Can Report: “Cost Free” Color Phenomenality Outside Focal Attention. *Psychological Science*, v. 25, n. 7, pp. 1394-1403, 2014.
- BRONFMAN, Z. Z. et al. The Transition to Minimal Consciousness through the Evolution of Associative Learning, *Frontiers in Psychology*, v. 7, pp. 1-16, 2016.
- BRONFMAN, Z. Z. et al. Impoverished or Rich Consciousness Outside Attentional Focus: Recent Data Tip the Balance for *Overflow*. *Mind and Language*, v. 34, n. 4, pp. 423-444, 2019.
- BROWN, R. et al. Understanding the High-Order Approach to Consciousness. *Trends in Cognitive Sciences*, v. 23, n. 9, pp. 754-768, 2019.
- BROWNING, H. The Measurability of Subjective Animal Welfare. *Journal of Consciousness Studies*, v. 29, n. 3-4, pp. 150-179, 2022.
- BROWNING, H.; VEIT, W. The Measurement Problem of Consciousness. *Philosophical Topics*, v. 48, n. 1, pp. 85-108, 2020.
- _____. The Sentience Shift in Animal Research. *The New Bioethics*, pp. 1-16, 2022.
- BROWNING, H.; BIRCH, J. Animal Sentience. *Philosophy Compass*, pp. 1-14, 2022.
- BURGE, T. Individualism and the Mental. *Midwest Studies in Philosophy*, v. 4, pp. 73-121, 1979.
- _____. *Origins of Objectivity*. New York: Oxford University Press, 2010.
- _____. *Perception: First Form of Mind*. New York: Oxford University Press, 2022.
- CALVO, P. What Is It Like to Be a Plant? *Journal of Consciousness Studies*, v. 24, n. 9-10, pp. 205-227, 2017.
- CARRUTHERS, P.; VEILLET, B. The Phenomenal Concept Strategy. *Journal of Consciousness Studies*, v. 14, n. 9-10, pp. 212-236, 2007.
- CARRUTHERS, P. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press, 2000.
- _____. Why the Question of Animal Consciousness Might Not Matter Very Much. *Philosophical Psychology*, v. 18, n. 1, pp. 83-102, 2005.
- _____. *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*. New York: Oxford University Press, 2006.
- _____. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press, 2011.

- _____. Evolution of Working Memory. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, suppl. 2, pp. 10371-10378, 2013.
- _____. *The Centered Mind: What the Science of Working Memory Show Us About the Nature of Human Thought*. Oxford: Oxford University Press, 2015.
- _____. Block's Overflow Argument. *Pacific Philosophical Quarterly*, v. 98, n. 1, pp. 65-70, 2017a.
- _____. In Defense of First-Order Representationalism. *Journal of Consciousness Studies*, v. 24, n. 5-6, pp. 74-87, 2017b.
- _____. The Problem of Animal Consciousness. *Proceedings and Addresses of the APA*, v. 92, pp. 179-205, 2018a.
- _____. Comparative Psychology Without Consciousness. *Consciousness and Cognition*, v. 63, pp. 47-60, 2018b.
- _____. *Human and Animals Minds: The Consciousness Questions Laid to Rest*. New York: Oxford University Press, 2019.
- _____. Stop Caring About Consciousness. *Philosophical Topics*, v. 48, n. 1, pp. 1-20, 2020a.
- _____. How Mindreading Might Mislead Cognitive Science. *Journal of Consciousness Studies*, v. 27, n. 7-8, pp. 195-219, 2020b.
- CHAI, W. J. et al. Working Memory from the Psychological and Neurosciences Perspectives: A Review. *Frontiers in Psychology*, v. 9, pp. 1-16, 2018.
- CHALMERS, D. J. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies*, v. 2, n. 3, pp. 200-219, 1995.
- _____. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press, 1996.
- _____. Availability: The Cognitive Basis of Experience? In: BLOCK, N. et al. (Orgs.) *The Nature of Consciousness: Philosophical Debates*. Cambridge, MA: MIT Press, 1997.
- _____. What Is a Neural Correlate of Consciousness? In: *The Character of Consciousness*. New York: Oxford University Press, 2010a.
- _____. Consciousness and Its Place in Nature. In: *The Character of Consciousness*. New York: Oxford University Press, 2010b.
- _____. Phenomenal Concepts and the Explanatory Gap. In: *The Character of Consciousness*. New York: Oxford University Press, 2010c.
- _____. The Content of Phenomenal Concepts. In: *The Character of Consciousness*. New York: Oxford University Press, 2010d.

- _____. Panpsychism and Panprotopsychism. In: BRÜNTRUP, G.; JASKOLLA, L. (Orgs.) *Panpsychism: Contemporary Perspectives*. New York: Oxford University Press, 2017.
- _____. *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: W. W. Norton & Company, 2022.
- CHAUDHARY, U. et al. Spelling Interface Using Intracortical Signals in a Completely Locked-In Patient Enabled via Auditory Neurofeedback Training. *Nature Communications*, v. 13, pp. 1-9, 2022.
- CHITTKA, L. *The Mind of a Bee*. New Jersey: Princeton University Press, 2022.
- CHUDEK, M. et al. Do Minds Switch Bodies? Dualist Interpretations Across Ages and Societies. *Religion, Brain & Behavior*, v. 8, n. 4, pp. 354-368, 2018.
- CHURCHLAND, P. S. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: The MIT Press, 1986.
- _____. *Touching a Nerve: The Self as Brain*. New York: W. W. Norton & Company, 2013.
- _____. What is Neurophilosophy and How Did Neurophilosophy Get Started? *Journal of NeuroPhilosophy*, v. 1, n. 1, pp. 1-16, 2022.
- CLARK, R. E.; SQUIRE, L. R. Classical Conditioning and Brain Systems: The Role of Awareness. *Science*, v. 280, n. 77, pp. 77-81, 1998.
- _____. Human Eyeblink Classical Conditioning: Effects of Manipulating Awareness of the Stimulus Contingencies. *Psychological Science*, v. 10, n. 1, pp. 14-18, 1999.
- CLARK, R. E. et al. Trace and Delay Eyeblink Conditioning: Contrasting Phenomena of Declarative and Nondeclarative Memory. *Psychological Science*, v. 12, n. 4, pp. 304-308, 2001.
- CLARK, R. E. et al. Classical Conditioning, Awareness, and Brain Systems. *Trends in Cognitive Sciences*, v. 6, n. 12, pp. 524-531, 2002.
- COHEN, M. A. et al. Perceptual Awareness Occurs Along a Graded Continuum: Evidence from Psychophysical Scaling. *Não publicado*, 2022. [Acessado em 29 de agosto de 2022] <https://psyarxiv.com/gsq7y/>
- COHEN, M. A.; DENNETT, D. C. Consciousness Cannot Be Separated from Function. *Trends in Cognitive Sciences*, v. 15, n. 8, pp. 358-364, 2011.
- COHEN, M. A.; DENNETT, D. C.; KANWISHER, N. What is Bandwidth of Perceptual Experience? *Trends in Cognitive Sciences*, v. 20, n. 5, pp. 324-335, 2016.
- CONNORS, M. H.; HALLIGAN, P. W. A Cognitive Account of Belief: A Tentative Road Map. *Frontiers in Psychology*, v. 5, pp. 1-14, 2015.

- COVA, F. et al. Is the Phenomenological Overflow Argument Really Supported by Subjective Reports? *Mind and Language*, pp. 1-29, 2020.
- COWAN, N. The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity, *Behavioral and Brain Sciences*, v. 24, n. 1, pp. 87-114, 2000.
- _____. What are the Differences Between Long-Term, Short-Term, and Working Memory? *Progress in Brain Research*, v. 169, pp. 323-338, 2008.
- _____. The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, v. 19, n. 1, pp. 51-57, 2010.
- COYNE, J. A. *Por que a Evolução é uma Verdade?* São Paulo: JSN Editora, 2014.
- CRICK, F.; KOCH, C. Towards a Neurobiological Theory of Consciousness. *Seminars in the Neuroscience*, v. 2, pp. 263-275, 1990.
- CROOK, R. J. Behavioral and Neurophysiological Evidence Suggests Affective Pain Experience in Octopus. *iScience*, v. 24, n. 3, pp. 1-9, 2021.
- CRUMP, A. Animal Sentience Science and Policy. *Animal Sentience*, n. 31, pp. 1-3, 2022.
- CRUMP, A. et al. Sentience in Decapod Crustaceans: A General Framework and Review of the Evidence. *Animal Sentience*, n. 32, pp. 1-34, 2022.
- CRUMP, A.; BIRCH, J. Separating Conscious and Unconscious Perception in Animals. *Learning & Behavior*, v. 49, n. 5, pp. 347-348, 2021.
- _____. Animal Consciousness: The Interplay of Neural and Behavioural Evidence. *Journal of Consciousness Studies*, v. 29, n. 3-4, pp. 104-128, 2022.
- CURTIS, C. E.; D'ESPOSITO, M. Persistent Activity in the Prefrontal Cortex During Working Memory. *Trends in Cognitive Sciences*, v. 7, n. 9, pp. 415-423, 2003.
- DAMÁSIO, A. *A Estranha Ordem das Coisas: As Origens Biológicas dos Sentimentos e da Cultura*. São Paulo: Companhia das Letras, 2019.
- DAWKINS, M. S. Animal Welfare with and Without Consciousness. *Journal of Zoology*, v. 301, n. 1, pp. 1-10, 2017.
- _____. *The Science of Animal Welfare: Understanding What Animals Want*. New York: Oxford University Press, 2021.
- _____. The Science of Animal Sentience and the Politics of Animals Welfare Should Be Kept Separate. *Animal Sentience*, v. 31, pp. 1-4, 2022.
- DAWKINS, R. *O Relojoeiro Cego: A Teoria da Evolução Contra o Desígnio Divino*. São Paulo: Companhia das Letras, 2001.

- DEDIU, D.; LEVINSON, S. C. Neanderthal Language Revisited: Not Only Us. *Current Opinion in Behavioral Sciences*, v. 21, pp. 49-55, 2018.
- DEHAENE, S.; CHANGEUX, J-P. Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, v. 70, n. 2, pp. 200-227, 2011.
- DEHAENE, S. et al. A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks. *Proceedings of the National Academy of Sciences of the USA*, v. 95, pp. 14529-14534, 1998.
- DEHAENE, S. et al. Cerebral Mechanisms of Word Masking and Unconscious Repetition Priming. *Nature Neuroscience*, v. 4, pp. 752-758, 2001.
- DEHAENE, S. et al. What is Consciousness, and Could Machines Have It? *Science*, v. 358, n. 6362, pp. 486-492, 2017.
- DEHAENE, S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Penguin Books, 2014.
- _____. *How We Learn: Why Brains Learn Better Than Any Machine...for Now*. New York: Penguin Books, 2021.
- DENNETT, D. C. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster Paperbacks, 1995.
- _____. *From Bacteria to Bach and Back: The Evolution of Minds*. New York: W. W. Norton & Company, 2017.
- DENTON, D. A. et al. The Role of Primordial Emotions in the Evolutionary Origin of Consciousness. *Consciousness and Cognition*, v. 18, n. 2, pp. 500-514, 2009.
- DERRIEN, D. et al. The Nature of Blindsight: Implications for Current Theories of Consciousness. *Neuroscience of Consciousness*, v. 7, n. 2, pp. 1-14, 2022.
- DIAMOND, J. *O Terceiro Chimpanzé: A Evolução e o Futuro do Ser Humano*. 5 ed. Rio de Janeiro: Record, 2020.
- _____. *Armas, Germes e Aço*. 27 ed. Rio de Janeiro: Record, 2021.
- DOERIG, A. et al. Hard Criteria for Empirical Theories of Consciousness. *Cognitive Neuroscience*, v. 12, n. 2, pp. 41-62, 2021.
- DOESBURG, S. M.; WARD, L. M. Corticothalamic Necessity, Qualia, and Consciousness. *Behavioral and Brain Sciences*, v. 30, n. 1, pp. 90-91, 2007.
- DOMJAN, M. *The Principles of Learning*. 7 ed. Stamford: Wadsworth Publishing, 2014.
- DOUVEN, I. *Abduction*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2021 [Acessado em 22 de julho de 2021] <https://plato.stanford.edu/archives/sum2021/entries/abduction>

- DRETSKE, F. I. Perception and Other Minds. *Noûs*, v. 7, n. 1, pp. 34-44, 1973.
- _____. *Naturalizing the Mind*. Cambridge, MA: MIT Press, 1995.
- _____. The Mind's Awareness of Itself. *Philosophical Studies*, v. 95, pp. 103-124, 1999.
- DUARTE, M. L. et al. Hidranencefalia – Uma Doença Rara com Diagnóstico Ultrassonográfico. *SALUSVITA*, v. 38, n. 3, pp. 685-692, 2019.
- DUARTE, R. et al. Black Hole Weather Forecasting with Deep Learning: A Pilot Study. *Monthly Notices of the Royal Astronomical Society*, 2022.
- DUFOUR, V. et al. Corvids Can Decide if a Future Exchange is Worth Waiting for. *Biology Letters*, v. 8, n. 2, pp. 201-204, 2012.
- DUNG, L. Does Illusionism Imply Skepticism of Animal Consciousness? *Synthese*, v. 200, n. 238, pp. 1-19, 2022.
- EDELMAN, D. B. et al. Identifying Hallmarks of Consciousness in Non-mammalian Species. *Consciousness and Cognition*, v. 14, n. 1, pp. 169-187, 2005.
- EDELMAN, D. B.; SETH, A. K. Animal Consciousness: A Synthetic Approach. *Trends in Neuroscience*, v. 32, n. 9, pp. 476-484, 2009.
- EVANS, G. Can There Be Vague Objects? *Analysis*, v. 38, n. 4, p. 208, 1978.
- EYSENCK, M. W.; KEANE, M. T. *Cognitive Psychology: A Student's Guide*. 8 ed. New York: Routledge, 2020.
- FAZEKAS, P.; NEMETH, G. Multidimensional Degrees of Consciousness: Qualities of Experiences During Wakeful Perception, Working Memory Tasks, Mental Imagery, Mind Wandering and Dreaming. In: OVERGAARD, M. et al. (Orgs.) *Beyond Neural Correlates of Consciousness*. New York: Routledge, 2020.
- FAZEKAS, P.; OVERGAARD, M. Multidimensional Models of Degrees and Levels of Consciousness. *Trends in Cognitive Sciences*, v. 20, n. 10, pp. 715-716, 2016.
- FEINBERG, T.; MALLATT, J. The Evolutionary and Genetics Origins of Consciousness in the Cambrian Period Over 500 Million Years Ago. *Frontiers in Psychology*, v. 4, pp. 1-27, 2013.
- _____. *The Ancient Origins of Consciousness: How The Brain Created Experience*. Cambridge, MA: MIT Press, 2017.
- _____. *Consciousness Demystified*. Cambridge, MA: MIT Press, 2018.
- _____. Phenomenal Consciousness and Emergence: Eliminating the Explanatory Gap. *Frontiers in Psychology*, v. 11, pp. 1-15, 2020.

- FERREIRA, C. P. *The Development of Consciousness*. Tese (Doutorado em Filosofia) – Instituto de Filosofia e Ciências Sociais, Universidade Federal do Rio de Janeiro. Rio de Janeiro, p. 168, 2017.
- FINN, E. S. et al. Layer-Dependent Activity in Human Prefrontal Cortex During Working Memory. *Nature Neuroscience*, v. 22, pp. 1687-1695, 2019.
- FISCHER, B. *Animal Ethics: A Contemporary Introduction*. New York: Routledge, 2021.
- FITZPATRICK, S. Against Morgan's Canon. In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018.
- FLEMING, S. M. *Know Thyself: The New Science of Self-Awareness*. London: John Murray, 2021.
- FODOR, J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press, 1987.
- FRANCKEN, J. C. An Academy Survey on Theoretical Foundations, Common Assumptions and the Current State of Consciousness Science. *Neuroscience of Consciousness*, v. 2022, n. 1, pp. 1-13, 2022.
- FRANKISH, K. Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*, v. 23, n. 11-12, pp. 11-39, 2016.
- FUNASHI, S. Working Memory in the Prefrontal Cortex. *Brain Sciences*, v. 7, n. 5, pp. 1-22, 2017.
- FUTUYMA, D. J. *Biologia Evolutiva*. 3 ed. Ribeirão Preto: FUNPEC Editora, 2009.
- GALLAGHER, S. Direct Perception in the Intersubjective Context. *Consciousness and Cognition*, v. 17, n. 2, pp. 535-543, 2008.
- GALWAY-WITHAM, J.; STRINGER, C. How Did *Homo sapiens* Evolve? *Science*, v. 360, n. 6395, pp. 1296-1298, 2018.
- GARSON, J. Function and Teleology. In: SARKAR, S.; PLUTYNSKI, A. (Orgs.) *A Companion to the Philosophy of Biology*. Oxford: Blackwell Publishing, 2008.
- GAZZANIGA, M. S. *Psychological Science*. 6 ed. New York/London: W. W. Norton & Company, 2018.
- GIBBONS, M. et al. Motivational Trade-Offs and Modulation of Nociception in Bumblebees. *Proceedings of the National Academy of Sciences of the USA*, v. 119, n. 31, pp. 1-3, 2022.
- GINSBURG, S.; JABLONKA, E. *The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness*. Cambridge, MA: MIT Press, 2019.

- _____. *Picturing the Mind: Consciousness Through the Lens of Evolution*. Cambridge, MA: MIT Press, 2022.
- GODFREY-SMITH, P. *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press, 1998.
- _____. *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press, 2009.
- _____. Cephalopods and the Evolution of the Mind. *Pacific Conservation Biology*, v. 19, n. 1, pp. 4-9, 2013.
- _____. Mind and Matter. *The Journal of Philosophy*, v. 113, n. 10, pp. 481-506, 2016.
- _____. Animal Evolution and the Origins of Experience. In: SMITH, D. L. (Org.) *How Biology Shapes Philosophy: New Foundations for Naturalism*. Cambridge: Cambridge University Press, 2017.
- _____. Towers and Trees in Cognitive. In: HUEBNER, B. (Org.) *The Philosophy of Daniel Dennett*. New York: Oxford University Press, 2018a.
- _____. The Evolution of Consciousness in Phylogenetic Context. In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018b.
- _____. *Outras Mentas: O Polvo e a Origem da Consciência*. São Paulo: Todavia, 2019a.
- _____. Evolving Across the Explanatory Gap. *Philosophy, Theory, and Practice in Biology*, v. 11, n. 1, pp. 1-24, 2019b.
- _____. *Metazoa: Animal Life and the Birth of the Mind*. New York: Farrar, Straus and Giroux, 2020a.
- _____. Varieties of Subjectivity. *Philosophy of Science*, v. 87, n. 5, pp. 1150-1159, 2020b.
- _____. Gradualism and the Evolution of Experience. *Philosophical Topics*, v. 48, n. 1, pp. 201-220, 2020c.
- _____. *Theory and Reality: An Introduction to the Philosophy of Science*. 2 ed. Chicago: The University of Chicago Press, 2021.
- GOFF, P. *Consciousness and Fundamental Reality*. New York: Oxford University Press, 2017.
- GOFF, P. et al. *Panpsychism*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2022 [Acessado em 14 de maio de 2022] <https://plato.stanford.edu/archives/sum2022/entries/panpschysm/>
- GOLDMAN-RAKIC, P. S. Cellular Basis of Working Memory. *Neuron*, v. 14, pp. 477-485, 1995.

GOODALE, M. A. et al. Separate Neural Pathways for the Visual Analysis of Object Shape in Perception and Prehension. *Current Biology*, v. 4, n. 7, pp. 604-610, 1994.

GOODALE, M. A. How (and Why) the Visual Control of Action Differs from Visual Perception. *Proceedings of the Royal Society B*, v. 281, n. 1785, pp. 1-9, 2014.

GOODALE, M. A.; MILNER, A. D. Two Visual Pathways – Where Have They Taken Us and Where Will They Lead in Future? *Cortex*, v. 98, pp. 283-292, 2018.

GOULD, S. J.; LEWONTIN, R. C. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London, Series B*, v. 205, n. 1161, pp. 581-598, 1979.

GRATÃO, M. S. et al. Primeiros Bípedes. In: NEVES, W. A. et al. (Orgs.) *Assim Caminhou a Humanidade*. 3 ed. São Paulo: Palas Athena, 2022.

GRAU, J. W. Learning from the Spinal Cord: How the Study of Spinal Cord Plasticity Informs Our View of Learning. *Neurobiology of Learning and Memory*, v. 108, pp. 155-171, 2017.

GRAZIANO, M. S. A.; WEBB, T. W. From Sponge to Human: The Evolution of Consciousness. *Evolution of Nervous Systems (Second Edition)*, v. 3, pp. 547-554, 2017.

GRAZIANO, M. S. A. *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. New York: W. W. Norton & Company, 2019.

_____. A Conceptual Framework for Consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, v. 119, n. 18, pp. 1-9, 2022.

GREENE, B. *O Universo Elegante: Supercordas, Dimensões Ocultas e a Busca da Teoria Definitiva*. São Paulo: Companhia das Letras, 2001.

GRICE, H. P. Meaning. *The Philosophical Review*, v. 66, n. 3, pp. 377-388, 1957.

GRIFFIN, D. R.; SPECK, G. B. New Evidence of Animal Consciousness. *Animal Cognition*, v. 7, pp. 5-18, 2004.

GRIFFIN, D. R. *Animal Minds*. Chicago: The University of Chicago Press, 1992.

_____. Afterword: What Is It Like? In: BEKOFF, M.; ALLEN, C.; BURGHARDT, G. M. (Orgs) *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge, MA: MIT Press, 2002.

GUGER, C. et al. Complete Locked-In and Locked-In Patients: Command Following Assessment and Communication with Vibro-Tactile P300 and Motor Imagery Brain-Computer Interface Tools. *Frontiers in Neuroscience*, v. 11, pp. 1-11, 2017.

GÜNTÜRKÜN, O. The Avian ‘Prefrontal Cortex’ and Cognition. *Current Opinion in Neurobiology*, v. 15, n. 6, pp. 686-693, 2005.

- GÜNTÜRKÜN, O.; BUGNYAR, T. Cognition without Cortex. *Trends in Cognitive Sciences*, v. 20, n. 4, pp. 291-303, 2016.
- HA, H. et al. Functional Organization of a Neural Network for Aversive Olfactory Learning in *Caenorhabditis elegans*. *Neuron*, v. 68, n. 6, pp. 1173-1186, 2010.
- HAAG, E.; DYSON, K. Trade-off Between Safety and Feeding in the Sea Anemone *Anthopleura aureoradiata*. *New Zealand Journal of Marine and Freshwater Research*, v. 48, n. 4, pp. 540-546, 2014.
- HAECKEL, E. Our Monism: The Principles of a Consistent Unitary World-View. *The Monist*, v. 2, n. 4, pp. 481-486, 1892.
- HALINA, M. et al. Evolutionary Transition Markers and the Origins of Consciousness. *Journal of Consciousness Studies*, v. 29, n. 3-4, pp. 62-77, 2022.
- HALL, G. Is Consciousness Vague? *Australasian Journal of Philosophy*, 2022.
- HAMM, A. O. et al. Affective Blindsight: Intact Fear Conditioning to a Visual Cue in a Cortically Blind Patient. *Brain*, v. 126, n. 2, pp. 267-275, 2003.
- HANKE, F. D.; KELBER, A. The Eye of the Common Octopus (*Octopus vulgaris*). *Frontiers in Physiology*, v. 10, pp. 1-11, 2020.
- HARLEY, T. A. *The Science of Consciousness: Waking, Sleeping and Dreaming*. New York: Cambridge University Press, 2021.
- HENDERSON, L. Bayesianism and Inference to the Best Explanation. *The British Journal for the Philosophy of Science*, v. 65, n. 4, pp. 687-715, 2014.
- HENRICH, J. *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. New York: Picador, 2021.
- HILL, C. *Consciência*. São Paulo: Editora Unesp, 2010.
- HOCHNER, B. et al. The Octopus: A Model for a Comparative Analysis of the Evolution of Learning and Memory Mechanisms. *Biological Bulletin*, v. 210, n. 3, pp. 308-317, 2006.
- HOFFMANN, D. L.; et al. U-th Dating Carbonate Crusts Reveals Neandertal Origin of Iberian Cave Art. *Science*, v. 359, n. 6378, pp. 912-915, 2018.
- HUME, D. *Tratado da Natureza Humana*. 2 ed. São Paulo: Editora Unesp, 2009.
- IRVINE, E.; SPREKAK, M. Eliminativism About Consciousness. In: KRIEGEL, U. (Org.) *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press, 2020.
- JACK, A. I. et al. fMRI Reveals Reciprocal Inhibition Between Social and Physical Cognitive Domains. *NeuroImage*, v. 66, pp. 385-401, 2013.
- JACKSON, F. Epiphenomenal Qualia. *The Philosophical Quarterly*, v. 32, n. 127, pp. 127-136, 1982.

- _____. What Mary Didn't Know. *Journal of Philosophy*, v. 83, n. 5, pp. 291–295, 1986.
- JAMES, T. W. et al. Ventral Occipital Lesions Impair Object Recognition but Not Object-Directed Grasping: An fMRI Study. *Brain*, v. 126, n. 11, pp. 2463-2475, 2003.
- JAMIESON, D. Science, Knowledge, and Animal Minds. *Proceedings of the Aristotelian Society*, v. 98, pp. 79-102, 1998.
- KAAS, J. H. et al. Escaping the Nocturnal Bottleneck, and the Evolution of the Dorsal and Ventral Streams of Visual Processing in Primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 377, pp. 1-17, 2021.
- KAHNEMAN, D. *Rápido e Devagar: Duas Formas de Pensar*. Rio de Janeiro: Objetiva, 2012.
- KAHNEMAN, D. et al. *Ruído: Uma Falha no Julgamento Humano*. Rio de Janeiro: Objetiva, 2021.
- KAMIŃSKI, J. et al. Persistently Active Neurons in Human Medial Frontal and Medial Temporal Lobe Support Working Memory. *Nature Neuroscience*, v. 20, pp. 590-601, 2017.
- KANDEL, E. R. et al. *Princípios de Neurociências*. 5 ed. Porto Alegre: AMGH Editora, 2014.
- KANDEL, E. R. et al. *Principles of Neural Science*. 6 ed. New York: McGraw Hill Education, 2021.
- KAPOOR, V. et al. Decoding Internally Generated Transitions of Conscious Contents in the Prefrontal Cortex Without Subjective Reports. *Nature Communications*, v. 13, pp. 1-16, 2022.
- KAPPELER, P. M. *Animal Behaviour: An Evolutionary Perspective*. Cham: Springer, 2021.
- KATO, R. et al. Visual Instrumental Learning in Blindsight Monkeys. *Scientific Reports*, v. 11, 2021.
- KEY, B. Fish Do Not Feel Pain and its Implications for Understanding Phenomenal Consciousness. *Biology & Philosophy*, v. 30, n. 2, pp. 149-165, 2015.
- _____. Why Fish Do Not Feel Pain. *Animal Sentience*, n. 3, pp. 1-34, 2016.
- KEY, B. et al. Neural Design Principles for Subjective Experience: Implications for Insects. *Frontiers in Behavioral Neuroscience*, v. 15, pp. 1-20, 2021.
- KIM, C. et al. The Role of Frontopolar Cortex in Manipulation of Integrated Information in Working Memory. *Neuroscience Letters*, v. 595, pp. 25-29, 2015.
- KLEIN, C. Consciousness, Intention, and Command-Following in the Vegetative State. *The British Journal for the Philosophy of Science*, v. 68, n. 1, pp. 27-54, 2017.
- KLEIN, C.; BARRON, A. B. Insects Have the Capacity for Subjective Experience. *Animal Sentience*, n. 9, pp. 1-19, 2016.

KNOBE, J.; NICHOLS, S. *Experimental Philosophy*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2017 [Acessado em 06 de abril de 2022] <https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy>

KOCH, C. *The Quest for Consciousness: A Neurobiological Approach*. Englewood: Roberts & Company Publishers, 2004.

_____. *The Feeling of Life Itself: Why Consciousness is Widespread but Can't be Computed*. Cambridge, MA: MIT Press, 2020.

KORNBLITH, S. et al. Persistent Single-Neuron Activity During Working Memory in the Human Medial Temporal Lobe. *Current Biology*, v. 27, n. 7, pp. 1026-1032, 2017.

KORNBLITH, H. *Scientific Epistemology: An Introduction*. Oxford: Oxford University Press, 2021.

KOSGAARD, C. M. *Fellow Creatures: Our Obligations to the Other Animals*. New York: Oxford University Press, 2018.

KOUIDER, S. et al. How Rich is Consciousness? The Partial Awareness Hypothesis. *Trends in Cognitive Sciences*, v. 14, n. 7, pp. 301-307, 2010.

KOZUCH, B. Prefrontal Lesion Evidence Against Higher-Order Theories of Consciousness. *Philosophical Studies*, v. 167, pp. 721-746, 2014.

KRIPKE, S. *Naming and Necessity*. Malden, MA: Blackwell Publishing, 2018.

KRYPTOTOS, A-M. et al. Avoidance Learning: A Review of Theoretical Models and Recent Developments. *Frontiers in Behavioral Neuroscience*, v. 09, pp. 1-16, 2015.

LADYMAN, J.; ROSS, D. *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press, 2007.

LAGERCRANTZ, H.; CHANGEUX, J-P. The Emergence of Human Consciousness: From Fetal to Neonatal Life. *Pediatric Research*, v. 65, n. 3, pp. 255-260, 2009.

LAMME, V. A. F. Towards a True Neural Stance on Consciousness. *Trends in Cognitive Sciences*, v. 10, n. 11, pp. 494-501, 2006.

LANDMAN, R. et al. Large Capacity of Integrated Objects Before Change Blindness. *Vision Research*, v. 43, n. 2, pp. 149-164, 2003.

LAU, H. Theoretical Motivations for Investigating the Neural Correlates of Consciousness. *WIREs Cognitive Science*, v. 2, n. 1, pp. 1-7, 2011.

_____. *In Consciousness We Trust: The Cognitive Neuroscience of Subjective Experience*. Oxford: Oxford University Press, 2022.

- LAU, H.; ROSENTHAL, D. Empirical Support for High-Order Theories of Conscious Awareness. *Trends in Cognitive Sciences*, v. 15, n. 8, pp. 365-373, 2011.
- LEDOUX, J. E. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster, 1998.
- _____. Coming to Terms with Fear. *Proceedings of the National Academy of Science United States of America*, v. 111, n. 8, pp. 2871-2878, 2014.
- _____. *The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains*. New York: Viking, 2019.
- LEE, A. Y. Degrees of Consciousness. *Noûs*, pp. 1-23 2022.
- de LEÓN, M. S. P. et al. The Primitive Brain of Early Homo. *Science*, v. 372, n. 6538, pp. 165-171, 2021.
- LEÓN-CARRIÓN, J. et al. The Locked-In Syndrome: A Syndrome Looking for a Therapy. *Brain Injury*, v. 16, n. 7, pp. 571-582, 2002.
- LEVINE, J. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, v. 64, n. 4, pp. 354-361, 1983.
- LIPTON, P. *Inference to the Best Explanation*. 2 ed. New York: Routledge, 2004.
- LOGIE, R. H. et al. (Orgs.) *Working Memory: State of Science*. Oxford: Oxford University Press, 2021.
- LOUKOLA, O. J. et al. Bumblebees Show Cognitive Flexibility by Improving on an Observed Complex Behavior. *Science*, v. 355, n. 6327, pp. 833-836, 2017.
- LOW, P. et al. *The Cambridge Declaration on Consciousness*. 2012 [Acessado em 19 de outubro de 2021] <https://fcmconference.org/img/CambridgeDeclarationOnConsciousness.pdf>.
- LYON, P. The Cognitive Cell: Bacterial Behavior Reconsidered. *Frontiers in Microbiology*, v. 6, pp. 1-18, 2015.
- MACPHAIL, E. M. A Bird's Eye View of Consciousness. In: LILJENSTRÖM, H.; ÅRHEM, P. (Orgs.) *Consciousness Transitions: Phylogenetic, Ontogenetic, and Physiological Aspects*. Oxford: Elsevier, 2008.
- MALACH, R. Local Neuronal Relational Structures Underlying the Contents of Human Conscious Experience. *Neuroscience of Consciousness*, v. 7, n. 2, pp. 1-13, 2021.
- _____. The Role of the Prefrontal Cortex in Conscious Perception: The Localist Perspective. *Journal of Consciousness Studies*, v. 29, n. 7-8, pp. 93-114, 2022.
- MARSHALL, J. C.; HALLIGAN, P. W. Blindsight and Insight in Visuo-Spatial Neglect. *Nature*, v. 336, pp. 766-767, 1988.

- MARTÍ, A. P.; et al. The Symbolic Role of the Underground World Among Middle Paleolithic Neanderthals. *Proceedings of the National Academy of Sciences of the United States of America*, v. 118, n. 33, pp. 1-6, 2021.
- MASHOUR, G. A.; ALKIRE, M. T. Evolution of Consciousness: Phylogeny, Ontogeny, and Emergence from General Anesthesia. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, supl. 2, pp. 10357-10364, 2013.
- MASHOUR, G. A. et al. Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, v. 105, pp. 776-798, 2020.
- MATILLA, H. R. et al. Honey Bees (*Apis Cerana*) Use Animal Feces as a Tool to Defend Colonies Against Group Attack by Giant Hornets (*Vespa Soror*). *PLOS ONE*, v. 15, n. 12, 2020.
- MCBREARTY, S.; STRINGER, C. The Coast in Colour. *Nature*, v. 449, pp. 793-794, 2007.
- MCGINN, C. *The Mysterious Flame: Conscious Minds in a Material World*. New York: Basic Books, 1999.
- MCKENZIE, K. *Fundamentality and Grounding*. Cambridge: Cambridge University Press, 2022.
- MCLAUGHLIN, B.; BENNETT, K. *Supervenience*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2018 [Acessado em 13 de abril de 2022] <https://plato.stanford.edu/archives/sum2021/entries/supervenience>
- MEDEIROS, S. L. S. et al. Cyclic Alternation of Quiet and Active Sleep States in Octopus. *iScience*, v. 24, n. 4, pp. 1-19, 2021.
- MENZEL, R. et al. Honey Bees Navigate According to a Map-Like Spatial Memory. *Proceedings of the National Academy of Science United States of America*, v. 102, n. 8, pp. 3040-3045, 2005.
- MERKER, B. The Liabilities of Mobility: A Selection Pressure for the Transition to Consciousness in Animal Evolution. *Consciousness and Cognition*, v. 14, n. 1, pp. 89-114, 2005.
- _____. Consciousness Without a Cerebral Cortex: A Challenge for Neuroscience and Medicine. *Behavioral and Brain Sciences*, v. 30, n. 1, pp. 63-81, 2007.
- MICHEL, M; LAU, H. On Dangers Conflating Strong and Weak Versions of a Theory of Consciousness. *Philosophy and the Mind Sciences*, v.1, n. 2, pp. 1-13, 2020.
- MICHEL, M.; MORALES, J. Minority Reports: Consciousness and the Prefrontal Cortex. *Mind and Language*, v. 35, n. 4, pp. 493-513, 2020.

- MICHEL, M.; DOERIG, A.; A New Empirical Challenge for Local Theories of Consciousness. *Mind and Language*, 2021.
- MICHEL, M. Consciousness Science Underdetermined: A Short History of Endless Debates. *Ergo*, v. 6, n. 8, pp. 771-809, 2020.
- _____. Book Review: Human and Animal Minds: The Consciousness Laid to Rest. *The Philosophical Review*, v. 130, n. 4, pp. 619-623, 2021.
- _____. Conscious Perception and the Prefrontal Cortex: A Review. *Journal of Consciousness Studies*, v. 29, n. 7-8, pp. 115-157, 2022.
- MILLER, A. *Contemporary Metaethics: An Introduction*. 2 ed. Cambridge: Polity Press, 2013.
- MILNER, A. D. et al. Perception and Action in 'Visual Form Agnosia'. *Brain*, v. 114, n. 1, pp. 405-428, 1991.
- MILNER, A. D.; GOODALE, M. A. *The Visual Brain in Action*. 2 ed. New York: Oxford University Press, 2006.
- MORALES, J.; LAU, H. The Neural Correlates of Consciousness. In: KRIEGEL, U. (Org.) *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press, 2020.
- MORGAN, C. L. *An Introduction to Comparative Psychology*. 2 ed. London e Newcastle: The Walter Scott Publishing CO, 1903.
- MUDRIK, L. et al. Information Integration Without Awareness. *Trends in Cognitive Sciences*, v. 18, n. 9, pp. 488-496, 2014.
- MYLOPOULOS, M. Neurobiological Theories of Consciousness. In: YOUNG, B. D.; JENNINGS, C. D. (Orgs.) *Mind, Cognition, and Neuroscience: A Philosophical Introduction*. New York: Routledge, 2022.
- NACCACHE, L. Why and How Access Consciousness Can Account for Phenomenal Consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1755, pp. 1-9, 2018.
- NACCACHE, L. et al. Why Intracranial Electrical Stimulation of the Human Brain Suggests an Essential Role for Prefrontal Cortex in Conscious Processing: A Commentary on Raccah et al. *OSF Preprints*, 2021. [Acessado em 10 de junho de 2022] <https://osf.io/zrqp8>
- NAGEL, T. What is Like to Be a Bat? *The Philosophical Review*, v. 83, n. 4, pp. 435-50, 1974.
- NAVARRO-CEBRIÁN, A.; FISCHER, J. Precise Functional Connections Between the Dorsal Anterior Cortex and Areas Recruited for Physical Inference. *European Journal of Neuroscience*, v. 56, n. 1, pp. 3660-3673, 2022.

- NEANDER, K. The Teleological Notion of “Function”. *Australasian Journal of Philosophy*, v. 69, n. 4, pp. 454-468, 1991.
- NETO, C. M. et al. Origem e Dispersão do Gênero *Homo*. In: NEVES, W. A. et al. (Orgs.) *Assim Caminhou a Humanidade*. 3 ed. São Paulo: Palas Athena, 2022.
- NEWTON, I. *Principia: Princípios Matemáticos de Filosofia Natural*. São Paulo: Editora da Universidade de São Paulo, 2012.
- NIEDER, A. et al. A Neural Correlate of Sensory Consciousness in a Corvid Bird. *Science*, v. 369, n. 6511, pp. 1626-1629, 2020.
- NIEDER, A. Consciousness Without Cortex. *Current Opinion in Neurobiology*, v. 71, pp. 69-76, 2021.
- NILSSON, D-E. Eye Evolution and its Functional Basis. *Visual Neuroscience*, v. 30, pp. 5-20, 2013.
- NISBETT, R. E.; WILSON, T. D. Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, v. 84, n. 3, pp. 231-259, 1977.
- OKASHA, S. The Levels of Selection Debate: Philosophical Issues. *Philosophy Compass*, v. 1, n. 1, pp. 74-85, 2006.
- _____. *Agents and Goals in Evolution*. Oxford: Oxford University Press, 2020.
- OVERGAARD, M. Phenomenal Consciousness and Cognitive Access. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1755, pp. 1-6, 2018.
- _____. Insect Consciousness. *Frontiers in Behavioral Neuroscience*, v. 15, pp. 1-6, 2021.
- OWEN, A. M. et al. Detecting Awareness in the Vegetative State. *Science*, v. 313, n. 5792, p. 1402, 2006.
- PALMER, T. D.; RAMSEY, A. K. The Function of Consciousness in Multisensory Integration. *Cognition*, v. 125, pp. 353-364, 2012.
- PAPINEAU, D. *Thinking about Consciousness*. Oxford: Oxford University Press, 2002.
- _____. The Problem of Consciousness. In: KRIEGEL, U. (Org.) *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press, 2020.
- _____. *The Metaphysics of Sensory Experience*. Oxford: Oxford University Press, 2021.
- PARINS-FUKUSHI, C. et al. Phylogeny, Ancestors, and Anagenesis in the Hominin Fossil Record. *Paleobiology*, v. 45, n. 2, pp. 378-393, 2019.
- PASSINGHAM, R. E.; ROWE, J. B. *A Short Guide to Brain Imaging: The Neuroscience of Human Cognition*. Oxford: Oxford University Press, 2016.
- PASSINGHAM, R. E.; WISE, S. P. *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight*. Oxford: Oxford University Press, 2012.
- PASSINGHAM, R. E. *Understanding the Prefrontal Cortex: Selective Advantage, Connectivity, and Neural Operations*. Oxford: Oxford University Press, 2021.

- PAULINO, C. D. et al. *Estatística Bayesiana*. 2 ed. Lisboa: Fundação Calouste Gulbenkian, 2018.
- PAUTZ, A. *Perception*. New York: Routledge, 2021.
- _____. An Argument Against Papineau's Qualitative View of Sensory Experience. In: KRIEGEL, K. (Org.) *Oxford Studies in the Philosophy of Mind*. Oxford: Oxford University Press, no prelo. [Acessado em 19 de junho de 2022.] <https://philpapers.org/archive/PAUAAA-11.pdf>
- PEREIRA, R. H. S.; SOUZA FILHO, S. F. BOOK REVIEW: PAPINEAU, D. *The Metaphysics of Sensory Experience*. *Manuscrito*, v. 45, n. 2, pp. 267-276, 2022.
- PEREIRA, R. H. S. et al. Por Que Somos o Nosso Cérebro: O Enativismo Posto em Questão. *Trans/Form/Ação*, no prelo.
- PEREIRA, R. H. S. O Argumento do Conhecimento: Uma Nova Estratégia Físicista. *Perspectiva Filosófica*, v. 46, n. 1, pp. 37-52, 2019.
- _____. Nonconceptualism as the Key to a New Physicalist Strategy in the Knowledge Argument. *Humanities & Social Sciences Communications*, v. 7, pp. 1-7, 2020.
- _____. Um Modelo Alternativo para a Cognição Direta de Estados Mentais Elementares de Terceiros. *Revista de Filosofia Moderna e Contemporânea*, v. 9, n. 1, pp. 15-28, 2021.
- PERRY, J. *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press, 2001.
- PERRY, C. J.; BARRON, A. B. Honey Bees Selectively Avoid Difficult Choices. *Proceedings of the National Academy of Science United States of America*, v. 110, n. 47, pp. 19155-19159, 2013.
- PERRY, C. J. et al. Invertebrate Learning and Cognition: Relating Phenomena to Neural Substrate. *WIREs Cognitive Science*, v. 4, n. 5, pp. 561-582, 2013.
- PHILLIPS, I. B. Perception and Iconic Memory: What Sperling Doesn't Show. *Mind and Language*, v. 26, n. 4, pp. 381-411, 2011.
- _____. The Methodological Puzzle of Phenomenal Consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 373, n. 1755, pp. 1-9, 2018.
- PINKER, S. *Rationality: What it is, Why it Seems Scarce, Why it Matters*. New York: Viking, 2021.
- PIKA, S. et al. Ravens Parallel Great Apes in Physical and Social Cognitive Skills. *Scientific Reports*, v. 10, pp. 1-19, 2020.
- PLANER, R. J.; STERELNY, K. *From Signal to Symbol: The Evolution of Language*. Cambridge, MA: The MIT Press, 2021.
- PRINZ, J. *The Conscious Brain: How Attention Engenders Experience*. New York: Oxford University Press, 2012.

- _____. Attention, Working Memory, and Animal Consciousness. In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018.
- PROUST, J. Nonhuman Metacognition. In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018.
- PRÜFER, K. et al. The Bonobo Genome Compared with Chimpanzee and Human Genomes. *Nature*, v. 486, pp. 527-531, 2012.
- QUACH, K. T.; CHALASANI, S. H. Flexible Reprogramming of *Pristionchus pacificus* Motivation for Attacking *Caenorhabditis elegans* in Predator-Prey Competition. *Current Biology*, v. 32, pp. 1-14, 2022.
- RABY, C. R. et al. Planning for the Future by Western Scrub-Jays. *Nature*, v. 445, pp. 919-921, 2007.
- RACCAH, O.; BLOCK, N.; FOX, K. C. R. Does the Prefrontal Cortex Play an Essential Role in Consciousness? Insights from Intracranial Electrical Stimulation of the Human Brain. *The Journal of Neuroscience*, v. 41, n. 10, pp. 2076-2087, 2021.
- RAMSEY, W. M. *Representation Reconsidered*. Cambridge: Cambridge University Press, 2007.
- REBER, A. S. et al. All Living Organisms are Sentient. *Animal Sentience*, n. 31, pp. 1-3, 2022.
- REBER, A. S. Caterpillars, Consciousness and the Origins of Mind. *Animal Sentience*, n. 11, pp. 1-13, 2016.
- _____. *The First Minds: Caterpillars, 'Karyotes, and Consciousness*. Oxford: Oxford University Press, 2018a.
- _____. Sentient Plants? Nervous Mind? *Animal Sentience*, n. 11, pp. 1-5, 2018b.
- REINO UNIDO. Act Chapter 22, de 28 de abril de 2022. Animal Welfare (Sentience) Act 2022. *UK Public General Acts*. 28 abril 2022. [Acessado em 28 de julho de 2022] <https://www.legislation.gov.uk/ukpga/2022/22/enacted>
- RIDLEY, M. *Evolução*. Porto Alegre: Artmed, 2013.
- ROBINSON, W. *Epiphenomenalism*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2019 [Acessado em 21 de setembro de 2021] <https://plato.stanford.edu/archives/sum2019/entries/epiphenomenalism>
- RÖBLER, D. C. et al. Regularly Occurring Bouts of Retinal Movements Suggest an REM Sleep-Like State in Jumping Spiders. *Proceedings of the National Academy of Science United States of America*, v. 119, n. 33, pp. 1-3, 2022.
- ROSENTHAL, D. M. *Consciousness and Mind*. New York: Oxford University Press, 2005.

- ROWAN, A. N. et al. Animal Sentience: History, Science, and Politics. *Animal Sentience*, n. 31, pp. 1-20, 2022.
- SAPOLSKY, R. M. *Behave: The Biology of Humans at Our Best and Worst*. New York: Penguin Press, 2017.
- SCHLAG, J. Should the Superficial Superior Colliculus be part of Merker's Mesodiencephalic System? *Behavioral and Brain Sciences*, v. 30, n. 1, pp. 105-106, 2007.
- SCHLIESSER, E. Synthetic Philosophy. *Biology and Philosophy*, v. 34, n. 19, pp. 1-9, 2019.
- SCHNELL, A. K. et al. How Intelligent is a Cephalopod? Lessons from Comparative Cognition. *Biological Review*, v. 96, n. 1, pp. 162-178, 2020.
- SCHNELL, A. K. et al. Cuttlefish Exert Self-Control in a Delay of Gratification Task. *Proceedings of The Royal Society B*, v. 288, n. 1946, pp. 1-9, 2021.
- SCHUPBACH, J. N. *Bayesianism and Scientific Reasoning*. Cambridge: Cambridge University Press, 2022.
- SCHWITZGEBEL, E. Phenomenal Consciousness, Defined and Defended as Innocently as I Can Manage. *Journal of Consciousness Studies*, v. 23, n. 11-12, p. 224-235, 2016.
- _____. Bordeline Consciousness, When It's Neither Determinally True nor Determinally False That Experience Is Present. *Não publicado*, 2021. [Acessado em 30 de junho de 2022] <http://www.faculty.ucr.edu/~eschwitz/SchwitzPapers/BorderlineConsciousness-211220.pdf>
- SEARLE, J. R. *A Redescoberta da Mente*. 2 ed. São Paulo: Martin Fontes, 2006.
- _____. As Mentes dos Animais. In: SEARLE, J. *Consciência e Linguagem*. São Paulo: Martins Fontes, 2010.
- SEGUNDO-ORTIN, M.; CALVO, P. Consciousness and Cognition in Plants. *WIREs Cognitive Science*, v. 13, n. 2, pp. 1-23, 2022.
- SEITZ, R. J.; ANGEL, H-F. Belief Formation – A Driving Force for Brain Evolution. *Brain and Cognition*, v. 140, pp. 1-8, 2020.
- SERGENT, C. et al. Cueing Attention After the Stimulus is Gone Can Retrospectively Trigger Conscious Perception. *Current Biology*, v. 23, pp. 150-155, 2013.
- SETH, A. K. et al. Criteria for Consciousness in Humans and Other Mammals. *Consciousness and Cognition*, v. 14, n. 1, pp. 119-139, 2005.
- SETH, A. K. et al. Measuring Consciousness: Relating Behavioural and Neurophysiological Approaches. *Trends in Cognitive Sciences*, v. 12, n. 8, pp. 314-321, 2008.
- SETH, A. K.; BAYNE, T. Theories of Consciousness. *Nature Reviews Neuroscience*, pp. 1-14, 2022.
- SETH, A. K. *Being You: A New Science of Consciousness*. New York: Dutton, 2021.

- SHAPIRO, L.; SPAULDING, S. Embodied Cognition. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2021 [Acessado em 03 de julho de 2022] <https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/>
- SHEA, N.; BAYNE, T. The Vegetative State and the Science of Consciousness. *The British Journal for the Philosophy of Science*, v. 61, n. 3, pp. 459-484, 2010.
- SHEA, N.; HEYES, C. Metamemory as Evidence of Animal Consciousness: The Type That Does The Trick. *Biology and Philosophy*, v. 25, pp. 95-110, 2010.
- SHEA, N. Methodological Encounters with the Phenomenal Kind. *Philosophy and Phenomenological Research*, v. 84, n. 2, pp. 307-344, 2012.
- SHETTLEWORTH, S. J. *Cognition, Evolution, and Behavior*. Oxford: Oxford University Press, 2010.
- _____. Animal Cognition. In: FRANKISH, K.; RAMSEY, W. M. (Orgs.) *The Cambridge Handbook of Cognitive Science*. Cambridge: Cambridge University Press, 2012.
- SHEVLIN, H. Current Controversies in the Cognitive Science of Short-Term Memory. *Current Opinion in Behavioral Sciences*, v. 32, pp. 148-154, 2020.
- _____. Non-human Consciousness and the Specificity Problem: A Modest Theoretical Proposal. *Mind and Language*, pp. 1-18, 2021.
- SINGER, P. *Libertação Animal*. São Paulo: WMF Martins Fontes, 2010.
- SLIGTE, I. G. et al. V4 Activity Predicts the Strength of Visual Short-Term Memory Representations. *Journal of Neuroscience*, v. 29, n. 23, pp. 7432-7438, 2009.
- SMITH, J. D. et al. The High and Lows of Theoretical Interpretation in Animal-Metacognition Research. In: FLEMING, S. M.; FRITH, C. D. (Orgs.) *The Cognitive Neuroscience of Metacognition*. New York: Springer, 2014.
- SOBER, E. Evolution and the Problem of Others Minds. *The Journal of Philosophy*, v. 97, n. 7, pp. 365-386, 2000.
- _____. *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press, 2009.
- SOUZA FILHO, S. F. *Naturalising Intentionality: A Teleological Approach*. Tese (PhD em Filosofia) – Department of Philosophy, King's College London. London, p. 272, 2018.
- _____. A Dual Proposal of Minimal Conditions for Intentionality. *Synthese*, v. 200, n. 115, pp. 1-22, 2022.
- _____. Can Constancy Mechanisms Draw the Limits for Intentionality? *Disputatio: International Journal of Philosophy*, no prelo.
- SPERLING, G. The Information Available in Brief Visual Presentations. *Psychological Monographs: General and Applied*, v. 74, pp. 1-29, 1960.
- STACHO, M. et al. A Cortex-Like Canonical Circuit in the Avian Forebrain. *Science*, v. 369, n. 6511, 2020.
- STERELNY, K. From Hominins to Humans: How sapiens Became Behaviourally Modern.

- Philosophical Transactions of The Royal Society B*, v. 366, n. 1566, pp. 809-822, 2011.
- _____. *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, MA: MIT Press, 2012.
- _____. *The Pleistocene Social Contract: Culture and Cooperation in Human Evolution*. New York: Oxford University Press, 2021.
- STOLJAR, D. Physicalism and Phenomenal Concepts. *Mind & Language*, v. 20, n. 5, pp. 469-494, 2005.
- TAYLOR, D. E.; BURGESS, A. What in the World is Semantic Indeterminacy? *Analytic Philosophy*, v. 56, n. 4, pp. 298-317, 2015.
- TEEUWEN, R. R. M. et al. A Neuronal Basis of Iconic Memory in Macaque Primary Visual Cortex. *Current Biology*, v. 31, n. 24, pp. 5401-5414, 2021.
- THAGARD, P. *Bots and Beasts: What Makes Machines, Animals, and People Smart?* Cambridge, MA: MIT Press, 2021.
- THIBAUT, L. et al. Retrospective Attention Gates Discrete Conscious Access to Past Sensory Stimuli. *PLoS ONE*, v. 11, pp. 1-16, 2016.
- THOMPSON, E. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: The Belknap Press of Harvard University Press, 2007.
- _____. Could All Life Be Sentient? *Journal of Consciousness Studies*, v. 29, n. 3-4, pp. 229-265, 2022.
- TOMASELLO, M. *The Evolution of Agency: Behavioural Organization from Lizards to Humans*. Cambridge, MA: The MIT Press, 2022.
- TONONI, G.; KOCH, C. Consciousness: Here, There and Everywhere? *Philosophical Transactions of The Royal Society B*, v. 370, n. 1668, pp. 1-18, 2015.
- TRAVERS, E. et al. Learning Rapidly About the Relevance of Visual Cues Requires Conscious Awareness. *Quarterly Journal of Experimental Psychology*, v. 7, n. 8, pp. 1698-1713, 2018.
- TURNER-EVANS, D. B.; JAYARAMAN, V. The Insect Central Complex. *Current Biology*, v. 26, n. 11, pp. 453-457, 2016.
- TYE, M. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press, 1995.
- _____. Is Consciousness Vague or Arbitrary? *Philosophy and Phenomenological Research*, v. 56, n. 3, pp. 679-685, 1996.
- _____. *Consciousness Revisited: Materialism Without Phenomenal Concepts*. Cambridge, MA: MIT Press, 2009.
- _____. *Tense Bees and Shell-Shocked Crabs: Are Animals Conscious?* New York: Oxford University Press, 2017.
- _____. Do Fish Have Feelings? In: ANDREWS, K.; BECK, J. (Orgs.) *The Routledge Handbook of Philosophy of Animal Minds*. New York: Routledge, 2018.

- _____. *Vagueness and the Evolution of Consciousness: Through the Looking Glass*. New York: Oxford University Press, 2021.
- _____. From Electrons to Elephants: Context and Consciousness. In: WUPPULURI, S.; STEWART, I. (Orgs.) *From Electrons to Elephants and Elections: Exploring the Role of Content and Context*. Cham: Springer, 2022.
- VAN GULICK, R. *Consciousness*. In: ZALTA, E. (Org.) *The Stanford Encyclopedia of Philosophy*. 2018 [Acessado em 27 de abril de 2021] <https://plato.stanford.edu/archives/spr2018/entries/consciousness/>
- VEIT, W. Complexity and the Evolution of Consciousness. *Biological Theory*, pp. 1-16, 2022a.
- _____. The Origins of Consciousness or the War of the Five Dimensions. *Biological Theory*, pp. 1-16, 2022b.
- de WAAL, F. B. M.; ANDREWS, K. The Question of Animal Emotions. *Science*, v. 375, n. 6587, pp. 1351-1352, 2022.
- de WAAL, F. *Are We Smart Enough to Know How Smart Animals Are?* New York: W. W. Norton & Company, 2017.
- _____. *O Último Abraço da Matriarca: As Emoções dos Animais e o que Elas Revelam Sobre Nós*. Rio de Janeiro: Zahar, 2021.
- WARD, J. *The Student's Guide to Cognitive Neuroscience*. 4 ed. New York: Routledge, 2020.
- WATKINS, S.; REES, G. The Human Superior Colliculus: Neither Necessary, nor Sufficient for Consciousness. *Behavioral and Brain Sciences*, v. 30, n. 1, p. 108, 2007.
- WEILNHAMMER, et al. An Active Role of Inferior Frontal Cortex in Conscious Experience. *Current Biology*, v. 31, pp. 1-13, 2021.
- WEISKRANTZ, L. *Blindsight: A Case of Study and Implications*. New York: Oxford University Press, 1986.
- WILLIAMS, J. R. G. W. Ontic Vagueness and Metaphysical Indeterminacy. *Philosophy Compass*, v. 3, n. 4, pp. 763-788, 2008.
- WILLIAMSON, T. *The Philosophy of Philosophy*. 2 ed. Hoboken: Wiley Blackwell, 2021.
- WILSON, E. O. et al. Chemical Releasers of Necrophoric Behavior in Ants. *Psyche: A Journal of Entomology*, v. 65, pp. 108-114, 1958.
- WYNN, T.; COOLIDGE, F. L. The Implications of the Working Memory Model for the Evolution of Modern Cognition. *International Journal of Evolutionary Biology*, v. 2011, 2011.
- WYNNE, C. D. L.; UDELL, M. A. R. *Animal Cognition: Evolution, Behavior and Cognition*. 3 ed. London: Bloomsbury Publishing, 2020.

YEATES, J. W. Ascribing Sentience: Evidential and Ethical Considerations in Policymaking. *Animals*, v. 12, n. 15, pp. 1-14, 2022.

ZULLO, L.; HOCHNER, B. A New Perspective on the Organization of an Invertebrate Brain. *Communicative and Integrative Biology*, v. 4, n. 1, pp. 26-29, 2011.