



UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA DA UFBA

**SOBRE A POSSIBILIDADE DE UMA CIÊNCIA DA
CONSCIÊNCIA: UMA LEITURA CRÍTICA DE CHALMERS
E CARRUTHERS**

RICARDO ALMEIDA MOTA RIBEIRO

Dissertação apresentada como requisito
parcial para a obtenção do título de Mestre
em Filosofia

Orientador: Prof. Dr. Waldormiro J. da Silva
Filho

Salvador, março de 2016

SUMÁRIO

AGRADECIMENTOS.....	p.3
RESUMO.....	p.4
ABSTRACT.....	p.5
INTRODUÇÃO.....	p.6
CAPÍTULO 1 O MISTÉRIO DA CONSCIÊNCIA.....	p.16
1.1 O <i>HARD PROBLEM</i> DA CONSCIÊNCIA.....	p.16
1.1.1 Reconstrução das Teses misterianistas.....	p.21
1.1.2 Problemas nas Teses misterianistas.....	p.26
1.2.SENTIDO PSICOLÓGICO E SENTIDO FENOMÊNICO.....	p.32
CAPÍTULO 2 SOBREVENIÊNCIA E O PROBLEMA DA CAUSAÇÃO....	p.37
2.1 CONCEITOS DE SOBREVENIÊNCIA.....	p.37
2.2 ARGUMENTO BIDIMENSIONAL E PROPOSIÇÕES FENOMÊNICAS.....	p.43
2.3 REDUCIONISMO E PRINCÍPIO DO FECHAMENTO CAUSAL.....	p.47
2.4.CIÊNCIA NÃO REDUCIONISTA E ANTINOMIA CAUSAL.....	p.51
CAPÍTULO 3 POR UMA CIÊNCIA DA CONSCIÊNCIA.....	p.57
3.1 DISCORDÂNCIAS FUNDAMENTAIS.....	p.65
3.1.1 Posicionamento de Carruthers.....	p.65
3.2.2 Resposta de Chalmers.....	p.68
3.2 ALGUNS DADOS SOBRE A MEMÓRIA DE TRABALHO.....	p.73
CONSIDERAÇÕES FINAIS.....	p.78
REFERÊNCIAS.....	p.85

AGRADECIMENTOS

Agradeço, acima de tudo, ao cuidado de minha tia, Maria Angelita de Almeida Motta, que vem sendo uma grande parceira na minha vida após o falecimento de minha mãe. Esse papel é mais discreto e não acadêmico, mas sem a sua presença constante em todos os setores da minha vida dificilmente eu teria condições de encerrar uma empreitada como a dissertação.

Agradeço também à breve conversa com o filósofo David Chalmers, cujo projeto de investigação sobre a ciência da consciência ocupa crucial papel em minha reflexão. Ele me mostrou o equívoco de seguir a orientação inicial que desejei a imprimir à pesquisa; a partir dessa conversa, dei uma guinada no meu foco e pude realizar o trabalho que ora se apresenta.

Agradeço também, especialmente, à paciente e cuidadosa revisão do orientador Prof. Dr. Waldomiro José da Silva Filho, a sua ciosa e, não raro, extremamente paciente – mais do que o mereceria - atenção para limar os meus deslizes como orientando; assim como também admiro o rigor acadêmico que muito me honra em ter sido seu aluno na graduação, bolsista de iniciação científica e agora orientando de mestrado.

Agradeço também aos professores Dr. Marco Aurélio Oliveira da Silva e Prof. Dr. César Schirmer dos Santos pelas valiosas contribuições na qualificação da dissertação, assim como em sua presença na banca de defesa, especialmente à indicação da obra de Peter Carruthers, *The Centered Mind*.

Cumprimento também as leituras críticas dos meus colegas no grupo de pesquisas da UFBA, dirigido por Prof. Dr. Waldomiro José da Silva Filho. Essas leituras me apontaram falhas que de outro modo não haveria percebido. Incorporei as suas observações em diversos trechos do trabalho.

Finalmente, o meu agradecimento à Fundação de Amparo a Pesquisa do Estado da Bahia (FAPESB) por me haver possibilitado, através do fomento à pesquisa resultante na bolsa de mestrado, dedicar-me, por dois anos, exclusivamente a realização deste trabalho.

RESUMO

A presente dissertação visa a enfrentar o problema mente-corpo. Três são os aspectos relevantes para esse trabalho: (1) mostrar que o sistema dualista, notadamente a versão de Chalmers do dualismo de propriedades, atende aos requisitos de uma boa resposta ao problema (2) mostrar que a tese misterianista, segundo a qual a consciência não pode ser conhecida, é falsa (3) mostrar que, desde os princípios da coerência estrutural e invariância na organização funcional, é possível estipular um NCC (correlato neural para a consciência) válido, compatível com a pesquisa científica contemporânea sobre a memória de trabalho apresentada por Peter Carruthers em *The Centered Mind* (4) Refutar a estratégia dos conceitos fenomênicos tal como se apresenta na réplica de Carruthers a Chalmers. Com o objetivo de alcançar estes resultados, a dissertação desenvolve-se ao duplo critério de exegese crítica de textos e argumentação original, pervadindo temas correlatos tais como as noções de sobreveniência lógica e natural, reducionismos ontológico e epistemológico da consciência, o princípio do fechamento, dentre outros tópicos de interesse na seara da questão mente-corpo. Ao final, o trabalho encerra inventariando breve ponderação a respeito da compatibilidade geral do citado sistema dualista com a investigação desenvolvida sobre o papel da memória de trabalho.

Palavras-chave: problema mente-corpo, sobreveniência, dualismo de propriedades, memória de trabalho, consciência

ABSTRACT

In this thesis we shall address the mind-body problem. In our treatment of the issue three main aspects shall be considered: (1) that the dualist system, notably Chalmers' version of property dualism, meets the requirements of a good response to the problem (2) that the mysterianist thesis, according to which consciousness can not be known, is false (3) that, according to the principles of structural coherence and invariance in functional organization, it is possible to set a valid NCC (neural correlate of consciousness), compatible with contemporary scientific research on working-memory as introduced in *The Centered Mind* by Peter Carruthers; (4) furthermore, we shall refute the phenomenal concept strategy as presented in Carruthers' response to Chalmers. In order to achieve all that, a double critical exegesis will be used, taking into account both the original argument and the texts, and considering related issues such as the logical and natural supervenience, the ontological and epistemological reductionism of consciousness and the principle of closure, among other topics of interest to the mind-body issue. We shall conclude by elaborating briefly on the overall compatibility between the aforementioned dualist system and the research undertaken regarding the role of working-memory.

Palavras-chave: mind-body problem, supervenience, property dualism, working memory, consciousness

INTRODUÇÃO

Essa dissertação trata do seguinte problema: é possível uma ciência da consciência? São enfrentadas algumas objeções a idéia de que seja possível uma ciência da consciência após o que se sugere que o melhor horizonte para a solução do problema geral é o proposto por David Chalmers. Logo em seguida, apresenta-se alguns traços relevantes do sistema deste filósofo para caracterizar a sua abordagem do problema e encerra em uma discussão da questão da causalidade mental. Esta última refere-se a um ponto pouco explorado por Chalmers, mas bastante importante na garantia da racionalidade da ação humana e em uma ciência que almeje dar conta deste aspecto.

Um dos tópicos centrais da filosofia da mente é o chamado problema mente/corpo (MCGINN, 1996, pp. 17-23). Pode-se afirmar que ele aparece em virtude da assunção de duas crenças intuitivas, mas antinômicas entre si. São elas: a) A crença segundo a qual os processos mentais e os processos físicos têm predicados diversos; aos primeiros, se atribui o não-determinismo, o livre-arbítrio, a transparência; para os segundos, indicam-se predicados opostos tais como o determinismo, a ausência de acesso direto, a objetividade. b) A crença, amparada pelos progressos da ciência cognitiva nesta direção, de que os processos mentais são estreitamente correlacionados a processos físicos cerebrais, determinados em algum grau por estes últimos.

Como há forte razões a favor de ambas as crenças, não é possível decidir-se trivialmente pela solução mais óbvia do problema, a saber, rejeitar uma delas em prol da outra. Essa solução dividir-se-ia em dois tipos, conforme a crença que se rejeitasse: 1) Rejeitar a crença *a* permitiria considerar, de maneira trivial, a mente um ente físico, regido pelas mesmas leis que vigoram no âmbito da natureza. 2) Rejeitar a crença *b* permitiria considerar a mente um objeto não-físico, não-determinado causal ou nomologicamente pelas leis da física. Logo, concluiríamos que a mente é distinta da matéria, regida por leis próprias e predicados psicológicos, a exemplo do predicado da não-extensão da res cogitans na teoria cartesiana.

A primeira solução é uma típica solução monista materialista reducionista. Esta proposta aparece em graus de radicalidade variando desde as que preservam o vocabulário psicológico até as eliminativistas (CHURCHLAND, 2006, p.151). Muitas vezes esposada por pesquisadores das ciências cognitivas embora não unanimemente, ela é, todavia, feixe de controvérsias acaloradas na comunidade filosófica. Já a segunda solução é paradigmaticamente o dualismo de substância, ainda vigoroso na cosmovisão popular, mas com poucos defensores no cenário do debate contemporâneo em filosofia da mente. As duas soluções podem revelar-se legítimas, mas não o são de modo *trivial*, isto é, sem a análise filosófica compatível com a complexidade do problema. Em artigo crítico, Tyler Burge apresenta, sinteticamente, sua evolução ao longo do período de 1950 a 2000 (BURGE, 2007, p. 440).

A primeira solução do problema a alcançar larga acolhida na comunidade filosófica contemporânea foi o behaviorismo. Utilizando o princípio verificacionista, o behaviorista elimina os termos mentais como desprovidos de sentido ou os reinterpreta à luz do seguinte critério: estados mentais são estados comportamentais observáveis, logo termos mentais só tem sentido se são compreendidos como termos comportamentais.

Tipicamente, um behaviorista sustentaria a tese segundo a qual termos mentais funcionam como termos disposicionais tais como "solúvel". Ser solúvel indica uma disposição do objeto para se dissolver uma vez encontradas as condições apropriadas: um torrão de açúcar, colocado em uma xícara de chá quente, é *solúvel*. Expressões mentais como Maria "deseja um café" não se refeririam a alguma propriedade contida da "mente" de Maria, mas se refeririam, ao invés disso, ao conjunto de possíveis comportamentos exibidos por Maria, tais como 1) Se lhe fosse oferecido um café Maria aceitaria 2) Se no cardápio de um restaurante serve-se café e chá, ela escolheria café, e assim por diante. Estes comportamentos são o significado de "desejar" o café.

Uma das teses do behaviorismo é que estados mentais são disposições com uma multiplicidade de possíveis realizações práticas. Neste sentido, cada termo costumeiramente associado a um estado mental, interno, involucra um feixe de comportamentos possíveis; este conjunto de comportamentos constitui o próprio estado mental, eliminando-se deste modo a referência a algum objeto

psicológico. Como se pode ver, essa perspectiva é compatível com a ontologia materialista e escapa, portanto, da antinomia assinalada na introdução.

Todavia, segundo Burge, a rejeição crescente ao behaviorismo ganhou força já a partir dos anos 50 acumulando mais de uma década de críticas. Estas críticas não se originaram de um evento específico - um ataque intelectual decisivo e identificável unitariamente - mas da percepção gradativa de que a abordagem behaviorista era restritiva e teoricamente infrutífera.

Paul Churchland (1998, p. 50) resume certos argumentos contra o behaviorismo. Suponha-se o condicional (1) do exemplo de Maria, acima. Ele seria falso caso Maria estivesse com raiva de quem a oferecesse o café e isso lhe bastasse para ela não aceitar. Imagine que este seja seu marido, e na noite anterior eles tivessem brigado. O café poderia ser um gesto tímido de conciliação ao qual Maria recusaria. O condicional (2), por sua vez, seria falso caso o café fosse mais caro do que o chá e ela julgasse que não valeria a pena pagar a diferença.

Nota-se intuitivamente que o número destes falseadores dos condicionais pode ser estendido indefinidamente. Pode-se imaginar uma infinidade de situações onde os condicionais são falseáveis. Caso o gesto de (1) seja do filho e ela esteja irritada com ele, ou caso o café está frio, etc. Para cada "estado mental" abrir-se-ia um leque de múltiplas vias de realização comportamental, cada um dos quais dependeria da ausência de uma conjunção indefinida de falseadores. Desta forma seria impossível, na prática, definir o mental exclusivamente "em termos de circunstâncias e comportamentos publicamente observáveis."

Não obstante a falha do behaviorismo em dar conta de problemas como esse, outras teorias materialistas surgiram, notadamente a teoria da identidade. Place (1954), em seu artigo *"Is Consciousness a Brain Process?"*, começa afirmando que certos termos cognitivos como "saber", "acreditar", "entender", assim como termos volitivos tais quais "querer" ou "pretender" podem ser analisados de acordo com o behaviorismo. Porém, conceitos como experiência, sensação, imaginação mental exigem um âmbito interno. A existência deste âmbito, todavia, não implica em afirmar que estes conceitos se referem a alguma substância distinta da física. Ao contrário, Place afirma que é possível tratá-los

desde uma tese naturalista que apresentará um núcleo muito simples: estados mentais são estados físicos, numericamente iguais uns aos outros. Cada estado mental x é um estado físico y .

A primeira parte do artigo (PLACE, 1954, pp. 44-45) consiste em analisar o conceito de identidade. Antes de entrar no mérito, Place lhe dá um escopo relativamente modesto: em vez de mostrar que a hipótese da identidade entre estados mentais e físicos é logicamente necessária, ele pretende mostrar que ela é cientificamente viável apenas. Place admite a inexistência de contradição lógica na hipótese segundo a qual é falsa a identidade, mas pretende aduzir argumentos para mostrar que a primeira hipótese segundo a qual ela é verdadeira é mais razoável. Inicialmente, o autor diferencia dois tipos de significado associado a "é": o "é" de definição e de composição. Quando dizemos que um triângulo equilátero é o polígono que tem três ângulos iguais, estamos definindo-o. Já quando dizemos que uma nuvem é uma massa de água não estamos definindo o que é a nuvem, mas estamos mostrando a sua composição. A nuvem é *feita* por uma massa de água. No primeiro caso, a verdade é *a priori*. No segundo, não.

Enunciar a frase "A consciência é um estado cerebral" quer dizer "a consciência *consiste* de um estado cerebral". Analiticamente, não existe nada que torne contraditório dizer que a consciência não seja um estado cerebral. Poderíamos imaginar um ser consciente que fosse imaterial. Muito embora não sejamos um tal ser, a sua possibilidade não pode ser eliminada *a priori*. Um outro traço dessa afirmação é que, ao contrário de afirmações como "esta roupa é um conjunto de trapos", ela se aplica a um conjunto amplo de fenômenos. Há uma identidade ontológica entre consciência e estado cerebral, embora não se siga da identidade lógica das expressões.

Place fornece o exemplo de identidade semelhante ao que encontramos entre consciência e processo cerebral. "Relâmpagos são o movimento de cargas elétricas". Do mesmo modo que o escrutínio cuidadoso do relâmpago não revela que ele seja o movimento de cargas elétricas, também a análise de estados conscientes não revelam que eles sejam certas operações cerebrais.

Putnam (1975 in LUDLOW, p.97) coloca alguns desafios a essa teoria. A inspiração do seu trabalho tampouco é mostrar *a priori* que estados mentais não

podem ser estados cerebrais. Ao contrário, ele desenvolve uma hipótese empírica, cuja verificação apresenta como sendo tão difícil quanto a da teoria da identidade, mas cuja vantagem seria ser menos restrita do que esta última.

Putnam faz um experimento mental: vamos supor um autômato probabilístico. Esse autômato realizaria uma série de inputs sensoriais em outputs mecânicos a partir de descrições que são codificadas pela máquina. Um input sensorial abriria um leque de probabilidades para tipos de *output* motores. Obviamente, a realização física deste autômato poderia veicular uma multiplicidade de descrições. Uma descrição S é tal que o sistema se comporta – em sua codificação de *inputs* e respostas motoras – de maneira unitária; e assim com S1, S2, etc.

Putnam (1975, pp. 96-104) é que é desnecessário conhecer a estrutura física da máquina para conhecer sua organização funcional. É este aspecto que importa ao se tentar determinar como o autômato interpreta os inputs em outputs “comportamentais”. Com efeito, Putnam alega que os defensores da teoria da identidade precisariam especificar um estado físico-químico tal que todo organismo possuísse a) um cérebro capaz de estar naquele estado b) este cérebro estivesse nesse estado.

As implicações disto é que o defensor da teoria da identidade precisa mostrar que quando um organismo sente “dor” ele possui o mesmo estado fisiológico e químico no cérebro que um outro organismo. Muito embora não seja de todo incomum encontrar essas virtuais identidades entre organismos de evolução biológica parecida, a hipótese claudica ao se tornar mais ambiciosa e universal. Afinal, ela afirma uma constituição ontológica, como se pode depreender do argumento de Place. Já que ela é válida para qualquer estado mental, então *todo* estado mental precisa ser o mesmo estado cerebral. Se organismos diferentes tem a sensação de vermelho e neles diversas estruturas cerebrais ativam, então isto significa que a sensação de vermelho não consiste no estado cerebral. Como de fato essa variedade existe na natureza, a teoria é falsa.

Essas são algumas respostas polêmicas que mostram como o problema se desenvolve, o tipo de argumentação e temática. O volume de detalhes e posições na história do problema é vastíssimo e inabarcável. Por isso mesmo,

cumprir fazer uma delimitação do campo ao qual se ater, a fim de tratá-lo nos limites de uma dissertação.

Esta dissertação está organizada em 3 capítulos, nos quais se expõe, respectivamente, as objeções a possibilidade da resolução do mistério da consciência, alguns traços do sistema de David Chalmers (que nos parece melhor resolvê-lo) e, finalmente, uma discussão sobre a noção de causalidade mental, que não foi devidamente desenvolvida pelo último. No presente texto, estão contidos os capítulos 1 e 2 em sua versão final, ausente dela, pois, o 3.

O primeiro capítulo da dissertação diz respeito ao "mistério da consciência". Nele apresentamos os motivos pelos quais a consciência é considerada um mistério. A consciência é, por um lado, um dado imediato, transparente e comum a todas as pessoas. Por outro lado, as pesquisas mais avançadas no campo das ciências cognitivas não lograram fornecer uma teoria aceitável que a desvendasse. Mais ainda: a questão a respeito da natureza da consciência parece apresentar desafios indelindáveis ao filósofo.

Nesse capítulo, serão especialmente investigados os argumentos de Colin McGinn, divididos em três teses básicas. Iniciamos com a tese geral do fechamento:

(FC): Uma mente M é cognitivamente fechada com respeito a propriedade P (ou teoria T) se e somente se os procedimentos de formação de conceito à disposição de M não podem se estender a captar P (ou compreender T)

McGinn advoga uma limitação inerente a nossa faculdade de conhecer, que deixaria a propriedade psicofísica P , a propriedade do cérebro causadora da consciência, fora de nossas capacidades cognitivas. Alguns são os indícios que ele aventa, argumentativamente:

1) nossas mentes fazem parte do mundo natural e tem limitações análogas a que nossos sentidos possuem. Nossa capacidade visual, por exemplo, limita-se a um certo espectro de raios, assim como a capacidade auditiva limita-se a uma faixa de Herz. Certos animais são capazes de alcançar percepções que são fechadas permanentemente a nossa experiência. Do mesmo modo, os poderes cognitivos humanos também seriam condicionados pela sua biologia e história evolutiva, e não existiria nenhuma prova segura de que possamos conhecer qualquer objeto da realidade.

2) A "sintaxe" própria da inteligência é a combinação de elementos e a analítica conceitual. Essa sintaxe emula a sintaxe da linguagem e da visão: a linguagem opera por meio de combinação de elementos atômicos em estruturas mais complexas dotadas de sentido. Estas combinações são feitas tendo em vista as funções sintáticas, semânticas e pragmáticas do termo. Elas se coadunam também com a capacidade natural dos seres humanos a adquirirem linguagem, a qual é ativada no ambiente propício a isto. Do mesmo modo, a ciência física essencialmente diz respeito a ordenação de elementos no espaço muito básicos - como os átomos - até complexos orgânicos e inorgânicos, seguindo leis formuladas parcialmente em linguagem matemática. Este caráter de nossa cognição não encontra paralelo na consciência, que é imediata, inespacial e unidirecional. A consciência não é capaz de por a si mesma como seu objeto. Ela possui um 'unidirecionamento' diferenciando-se quanto a isso de faculdades psicológicas como a imaginação ou de faculdades sensoriais, as quais fornecem um mapa do mundo amplamente detalhado, sob múltiplas 'direções'. A consciência, ao contrário, apresenta-se como um dado unitário e simples. Uma vez que estados de segunda ordem conscientes não conseguem ter a consciência como seu objeto, então, presumivelmente, certas propriedades da consciência escapam ao nosso conhecimento introspectivo que, de outro modo, não poderiam ser conhecidas.

3) O fato de o conhecimento humano se dar, essencialmente, por impressões e extrapolação de impressões torna a questão filosófica sobre a consciência de difícil resposta. Ao contrário das partículas subatômicas, cuja estipulação teórica ainda guarda analogia com o que examinamos diretamente no nível macroscópico, o caso da consciência é totalmente singular.

Serão discutidas algumas respostas a estes pontos, como:

- i) É perfeitamente possível que tenhamos limitações cognitivas intrínsecas em face de diversos objetos de cognição possível. Porém, inexistem evidências suficientes para se pensar que este é o caso.
- ii) Não é preciso ter o conhecimento do físico e da consciência diretamente para se obter conhecimento da propriedade psicofísica que explica a emergência da segunda. Além disso, problemas matemáticos contraintuitivos que levaram séculos mostram que a combinação de simplicidade e aparente insolubilidade não é forte o suficiente para sustentar a tese de McGinn.
- iii) O conhecimento humano não está tão estritamente vinculado a essas condições.

O segundo capítulo consiste na apresentação do núcleo da teoria de Chalmers (1996; 2010) noção de sobreveniência, reducionismo, fechamento causal. Ele começa por um breve panorama da noção de “sobreveniência” na literatura, detendo-se especialmente na formulação de Jaegwon Kim, uma das mais influentes e densas.

A noção de "sobreveniência" origina-se com Moore e Hare, no contexto da reflexão moral e estética e, rapidamente, invade o terreno da filosofia da mente. Grosso modo, a sobreveniência é a noção segundo a qual um conjunto de propriedades-A co-varia com um conjunto de propriedades-B. Ela divide-se em forte e fraca (em Kim) e lógica e natural (em Chalmers).

Formulada em termos modais, sobreveniência lógica é aquela que é válida necessariamente para todos os mundos possíveis, ao passo que a sobreveniência natural só o é para um determinado mundo, ou seja, ela é condicionada às leis de funcionamento do mundo. O que Chalmers busca com essa noção é encontrar um lugar específico para a consciência. Segundo ele, as propriedades conscientes sobrevivem ao mundo físico naturalmente, e não logicamente.

Apresentadas essas noções, tocamos no argumento bidimensional. O argumento bidimensional tem, grosso modo, a seguinte estrutura.

1. $P \& \sim Q$ é concebível.
2. Se $P \& \sim Q$ é concebível, $P \& \sim Q$ é metafisicamente possível.
3. Se $P \& \sim Q$ é metafisicamente possível, o materialismo é falso.

4. Materialismo é falso.

Aqui P é o conjunto de todas as verdades microfísicas do mundo, Q uma certa verdade fenomênica, por exemplo Q : alguém possui consciência Q' : alguém instancia um tipo específico de propriedade fenomênica. Se Q , $P \& \sim Q$ significa que neste mundo possível não há ninguém que possua consciência. Se Q' , $P \& \sim Q'$ significa que neste mundo possível há alguém que não instancia um tipo de propriedade fenomênica.

Uma das vantagens explicativas do argumento bidimensional é dar um tratamento unificado para diversos argumentos antirreducionistas. Após indicar a força explicativa que o argumento bidimensional traz, o capítulo abre uma nova seção. Nesta seção será explicado o fechamento causal.

Existem dois modos de se reduzir as propriedades fenomênicas às físicas. Um deles é a identificação pura e simples entre umas e outras. Se as propriedades fenomênicas são físicas então o âmbito próprio do fenomênico desaparece. Essa identificação aparece, particularmente, na teoria da identidade de Place e Smart e no reducionismo eliminativista, que rejeita o arcabouço de vocabulário da psicologia vulgar. Outro modo de operar a redução é fazê-lo quanto a sua epistemologia, isto é, reduzir as leis que regem o fenômeno da consciência às leis que regem o fenômeno físico.

Essa tendência reducionista encontra guarida em um princípio chamado de fechamento causal. O fechamento causal dispõe, em resumo, que todo efeito físico possui uma causa que é necessariamente físico. Este princípio parece ser improblemático quando examinado à luz dos dados das ciências empíricas menos básicas - como a química e biologia. Todavia, parece ser um problema para a concepção não reducionista da mente - a de Chalmers.

Se o mundo natural é causalmente fechado, então a consciência não exerce qualquer poder causal sobre o mundo. Mas, a nossa imagem de nós

mesmos é solidária de uma autocompreensão como agentes racionais. E uma das noções mais importantes de agente racional é justamente ser capaz de referir-nos ao nexos entre estados mentais e ação.

A esta altura, advinha-se um problema no sistema de Chalmers, o qual admite o fechamento causal ao mesmo tempo que busca preservar o não-reducionismo. Mas antes de focar completamente neste tópico, faz-se mister abordar mais alguns pontos, tais como dois princípios fundamentais para a sua ciência da consciência. Neste passo, a dissertação encerra um primeiro grande arco que vai da contra-argumentação a McGinn até a apresentação dos princípios da ciência da consciência de Chalmers. (MCGINN, 1997, 1999) sustenta a tese segundo a qual a consciência não pode ser conhecida. Uma vez ultrapassados os argumentos que se destinam a mostrar porque a consciência não pode ser objeto de ciência, é preciso pavimentar o percurso, apresentando alguns pontos do sistema de Chalmers sinteticamente, até a sua proposta: a construção de uma ciência da consciência.

Feito isto, ainda resta, porém, uma pergunta: como a consciência pode ter uma 'função' no mundo da vida quando o fechamento causal retira dela qualquer efeito no mundo? Chalmers diz, explicitamente, em uma resposta à agressiva resenha que John Searle (1997, pp. 136-176) lhe fez no *Mystery of Consciousness* que essa não era uma questão que ele trataria. A pergunta tampouco recebe uma resposta negativa.

|

|

CAPÍTULO 1: O "MISTÉRIO" DA CONSCIENCIA

Na seara do problema mente/corpo há um tópico que tem sido normalmente considerado sua pedra de toque. É aquilo que já foi chamado, por mais de um filósofo contemporâneo, "o mistério da consciência". Em sua eloquente defesa da desmistificação da consciência, Denett (1991, pp. 21-27) afirma que ela é o último dos mistérios sobreviventes. Temas como o começo do universo, a vida e a reprodução, o espaço-tempo e a gravitação universal passaram de enigmas a questões cujo tratamento, embora incompleto, revela-se promissor. Para cada um deles existe um avanço identificável na literatura científica e um norte de pesquisa. Ainda que se possam esperar substanciais mudanças na compreensão futura destas difíceis perguntas, existe um inequívoco senso de progresso no domínio dos respectivos campos de estudo.

Já com a consciência não acontece o mesmo. Ela continua a desafiar nossos padrões de investigação de um modo desconcertante, a revelar-se envolta numa opaca nevoa de obscuridade teórica. Um mistério, como diz Searle, é algo que não se sabe como pensar. Um problema é algo que nós sabemos como pensar, mas para cujas questões, por ele suscitadas, ainda não temos respostas definitivas. Ainda que não tenhamos uma teoria física completa; ainda que haja lacunas significativas na teoria evolucionista; ainda que ignoremos todos os detalhes da genética molecular essas impressionantes construções científicas, reveladoras da evolução na compreensão dos antigos mistérios, mostram que entramos em um caminho. Sabemos separar questões estéreis de questões fecundas; e, mesmo se estivermos errados em relação a alguns respostas, sabemos onde procurar as respostas certas. Os mistérios tornaram-se problemas (DENNET, 1991, p. 22). A consciência, porém, resiste a se tornar um problema, mas permanece um mistério.

1.1 O *HARD PROBLEM* DA CONSCIÊNCIA

Uma das possíveis saídas diante do impasse é considerar o problema da consciência *insolúvel*. McGinn (1997, p.29) busca mostrar por que este

problema é insolúvel. Inicialmente, McGinn admite que o cérebro seja a base causal da produção da consciência, ainda que não saibamos como isto é possível. Intuitivamente notamos que há um descompasso categorial intuitivo entre as ondas eletromagnéticas que a bioquímica cerebral aponta como constituindo a própria atividade neural consciente e o fenômeno qualitativo, subjetivo e medianamente transparente da consciência (MCGINN, 1993 p. 27). Como se viu na introdução, a teoria reducionista materialista busca resolver este descompasso em favor de uma identificação da mente (e a consciência especificamente) com a matéria. De acordo com essa explicação, os estados da consciência simplesmente *seriam* estados físicos.

McGinn (1999, p.28-29) discorda deste tipo de teoria. Ele admite o descompasso acima apontado e sublinha que ele abre traz uma dificuldade de explicação real, pois cumpre haver uma teoria que unifique consciência e matéria, teoria esta capaz de mostrar o elo psicofísico necessário para explicar como a consciência sobrevém à matéria. Uma vez que o fato desta sobreveniência lhe parece amplamente embasado na evidência empírica, o que resta ao filósofo é explicar como essa relação acontece.

Ao tentar explicá-lo, isto nos levaria a um caso típico de perplexidade filosófica. As perplexidades filosóficas surgem quando o senso comum deve ser posto entre os parênteses da reflexão crítica de uma forma marcadamente radical. Ao contrário de questionamentos científicos, cuja linguagem é bastante segura, o problema filosófico furta-se ao mesmo tratamento: o próprio modo de enunciá-lo já é dificultoso, espinhoso. O problema filosófico trai, em sua própria formulação, a confusão conceitual em que aquele que o formula se encontra, ao passo que a questão científica é precisa e compreensível o suficiente para não ser um "mistério". Por isso mesmo, pode ser resolvida de maneira "positiva".

Dois são os pontos: investigar o que faz o mistério ser misterioso, e, ato contínuo, mostrar que o mistério não pode ser solucionado. O primeiro dado é evidenciado pela questão a respeito de algo ser "like to be a bat" (NAGEL, 1974, pp.435-438); o que faz um morcego ter a experiência consciente (inapreensível a nossa subjetividade) de morcego? Como esta experiência pode se ligar a base psicofísica que, segundo todas as evidências empíricas, permite o aparecimento do dado da consciência? Para McGinn isto é um mistério, uma

vez que a própria formulação da questão impõe uma enorme dificuldade conceitual como se verá.

Segundo seu argumento para esclarecer qual questão precisamente ele deseja enfocar, o autor rejeita inicialmente uma interpretação possível falaciosa. Ele reforça que não está questionando como seria a experiência *para* o morcego. Para que essa pergunta fizesse sentido seria preciso que o morcego possuísse um estado mental de segunda ordem. A experiência seria um objeto deste estado mental de segunda ordem. Assim, para o morcego a sua própria experiência interna seria de tal ou qual forma; e o estado mental segundo essa experiência é de tal ou qual forma seria o foco da questão. Humanos frequentemente possuem estados mentais deste tipo. Podemos refletir acerca de um sentimento ou pensamento que temos e esse pensamento de segunda ordem pode ser objeto de análise filosófica. A análise enfocaria como essa experiência de segunda ordem surge, qual a sua natureza, etc. Mas não é do que se trata na investigação, mas sim do que é a experiência do morcego *enquanto tal*, em seu caráter qualitativo particular de experiência primária, e como essa experiência articula-se com a base material que, segundo todas as evidências empíricas, é necessária para que ela surja.

Entre as soluções propostas historicamente, McGinn (1989, p.529) assinala dois tipos: as naturalistas e aquelas que invocam entidades sobrenaturais ou intervenções divinas. Exemplo da primeira são soluções "construtivas" tais como o funcionalismo. Tal perspectiva caracteriza-se por oferecer uma descrição do comportamento associada a respostas funcionais. Da segunda, são as soluções de Leibniz ou Descartes. Estas apelam a substâncias espirituais ou a uma harmonização pré-estabelecida de natureza divina.

McGinn oferece uma alternativa peculiar: advoga que a explicação seja meramente natural, ao mesmo tempo que ela é incognoscível para nós, seres humanos. Inicialmente, ele apresenta a ideia de um fechamento cognitivo. Assim define (MCGINN, 1987, p.529):

(FC): Uma mente M é cognitivamente fechada com respeito a propriedade P (ou teoria T) se e somente se os procedimentos de

formação de conceito à disposição de M não podem se estender a captar P (ou compreender T)

Logo, o que é fechado a um tipo de mente pode ser aberto a outro. O que a mente de um rato ou macaco não consegue compreender, por um traço constitutivo de sua estrutura e não acidentalmente, pode nos ser plenamente acessível.

De fato, existem limitações ao nosso conhecimento que sabemos com absoluta certeza que não são franqueáveis. Tanto o passado quanto o futuro fornecem exemplo desses oceanos de ignorância. Nenhum homem atual pode saber exatamente a quantidade de dinossauros da Terra ou a de formigas na África em 1620. Isto porque o passado não deixa todos os traços dele para o presente. A nossa capacidade de investigação do passado é limitado por aquilo que podemos obter de registro dele. E embora seja até (teoricamente) possível contar todas as formigas existentes na África hoje, as dificuldades práticas insolúveis além da evidente falta de propósito de uma tal operação, tornam este fato desconhecido para qualquer pessoa. (MCGINN, Colin, 1999, p.48).

Duas são as concepções antropológicas do lugar do conhecimento (MCGINN, 1999). Por um lado, pode-se conceber o conhecimento como o caso de uma excepcional sorte. De acordo com essa perspectiva, a mente deve estar em uma relação especial e raramente propícia com o mundo para captar suas propriedades. A versão mais radical desta perspectiva encontra-se na crítica cética a faculdade de conhecer. O caso cético é seu extremo limite: nós *já* estamos nessa relação, adverte o cético, o nosso azar cognitivo é permanente e inerente à nossa condição. Por outro lado, pode-se imaginar uma situação onde a mente é naturalmente conhecedora; estamos, assim, em uma progressão regular rumo a um conhecimento cada vez mais abrangente da realidade; este traço de otimismo cognitivo, diz McGinn, é herdeiro da concepção iluminista da Razão, de sua confiança nos poderes da explicação científica e é dominante na contemporaneidade

O fechamento cognitivo é parcialmente solidário à primeira perspectiva. Se a relação entre mente e mundo é de tal ordem que somente uma certa faixa de propriedades nos é acessível, então é possível haver, de fato, um

fechamento cognitivo absolutamente infranqueável. Para sustentar essa tese ele (MCGINN, 1999) aponta a um dado bastante aceitável: mentes são produtos biológicos como os corpos, adaptados a cumprir uma série de funções. Utiliza o exemplo da percepção sensível: certas espécies são capazes de perceber diferentes propriedades do mundo, ex. certos animais que percebem o ultravioleta no espectro cromático.

Esta característica da mente tem a ver com um conceito de inteligência em que cumprem importante papel as distinções entre espécies, a modularidade e a relação entre mente e mundo. Neste conceito, a inteligência não é simplesmente a capacidade de estocar memória e lidar heurísticamente com os dados estocados, uma capacidade que seria idêntica às espécies variando exclusivamente quanto ao grau (MCGINN, 1999. p.39), assim como um HD de 250 Gb é diferente de um 1000 vezes maior ou um processador atual de um processador de 40 anos atrás.

Ao contrário, a inteligência humana, sendo uma faculdade da mente que se constitui precisamente na relação especial entre mundo e atividade humana tem um nítido enraizamento na especificidade de nossa constituição biológica geral e em nossa história evolutiva. Deste modo, McGinn esclarece, se uma criança fosse levada a Marte, onde houvesse uma sociedade de ETs possuidores de uma outra linguagem com regras sintáticas completamente diferentes das nossas, provavelmente ela seria incapaz de desenvolver esta linguagem. A mesma dificuldade surgiria na hipótese de as leis físicas da ótica serem repentinamente alteradas; tal revolução acarretaria a impossibilidade de formar imagens adequadas do entorno.

McGinn (1999, p.44) extrapola o exemplo para pensar em uma mente humeana, ou seja, uma mente cujas ideias sempre fossem cópias de impressões. Frases que comumente formulamos a respeito do mundo, e que envolvem noções de continuidade dos eventos no tempo e no espaço teriam suas condições de verdade em xeque porque estes conceitos seriam apenas a tradução de hábitos. Estaríamos habituados, diz Hume em um exemplo clássico na história da filosofia, a enxergar a causalidade entre o impacto de duas esferas atribuindo o movimento da segunda (impactada) à primeira. Mas, ao não poder formar a impressão do nexos de causalidade, e sim tão somente a dos eventos

separados, nossas atribuições de causalidade se dariam por uma frágil inferência oriunda do costume de ver o movimento da segunda esfera acontecer após o da primeira.

É bem verdade que é altamente improvável que a nossa mente funcione como Hume teorizou (MCGINN, 1999, p.44). Ainda assim, poderíamos imaginar uma tal mente. Uma tal inteligência não poderia alcançar certas propriedades das teorias físicas do mundo que desenvolvemos - o conceito de partículas inobserváveis estaria definitivamente fora de seu alcance, uma vez que não há possível impressão do inobservável.

Poderíamos também imaginar uma mente que só tivesse os sentidos da visão e do olfato. Desta maneira, a formação do conceito de espaço seria bastante prejudicada. Podemos imaginar que um protoconceito de espaço é formado (os sons mais altos e mais baixos, assim como os cheiros mais intensos e frágeis poderiam levá-lo a alguma noção bastante distorcida do espaço), mas é notório que ela seria radicalmente empobrecida.

1.1.1 Reconstrução das Teses Misterianistas

Três são os pontos de prova: (i) existe alguma propriedade natural do cérebro que explica o aparecimento da consciência, (ii) nós somos fechados cognitivamente a respeito desta propriedade (iii) não existe problema filosófico mente-corpo. Destes três, (i) está fora de discussão por enquanto. Partimos de uma concepção naturalista do mundo físico e da mente, e (iii) é uma consequência de (i) e (ii). O que nos importa é avaliar os argumentos para (ii).

Inicialmente, admite-se que a evidência histórica de falha na busca de uma solução para o problema mente-corpo é sugestiva, mas não conclusiva. É perfeitamente possível, afinal, que o problema esteja aguardando uma melhor formulação que pode acontecer amanhã. Não são incomuns 'mistérios' cognitivos que foram resolvidos após uma intuição de um indivíduo genial. Portanto, a falha na busca dessa solução não é conclusiva.

Todavia, existem razões mais fortes para considerarmos o problema insolúvel. McGinn discrimina dois meios de ter acesso a propriedade *P*: a) alcançar *P* por meio da investigação da consciência diretamente b) alcançar *P* por meio do estudo do cérebro. Rapidamente, ele descarta a primeira opção: embora a consciência seja particularmente transparente e acessível, não podemos compreender *P* acumulando descrições acerca dos nossos estados mentais ou articulando nossas autopercepções. Uma vez que a relação é entre *P* e uma mente física, o que existe no campo da consciência por si só não pode revelar esta relação. Chamemos esse dilema de tese 1 do argumento de McGinn (T1MA)

Um dos predicados da consciência que reforça este dilema é que ela parece não poder ser explicada mereologicamente (MCGINN, 1997), isto é, por meio de propriedades mais básicas (as do cérebro) combinadas de acordo com leis gerais. As propriedades dos líquidos, ao contrário, são obtidas deste modo. Outro predicado é que a consciência teria uma espécie de simplicidade unidirecional, oposta a multiplicidade e riqueza que nos fornecem os cinco sentidos.

Essa é uma intuição bastante interessante de McGinn. A idéia é que nas operações dos sentidos é possível uma riqueza de perspectivas. Por exemplo, se estamos diante de uma maçã, nós podemos enxergá-la desde diversos ângulos e níveis de distância e proximidade. Também podemos nos concentrar em pontos diferentes de sua superfície. O mesmo acontece, em maior ou menos grau, com as outras faculdades sensoriais. No caso das representações mentais também: podemos imaginar a maçã de múltiplas formas (inclusive inexistentes na realidade concreta); podemos representar estados mentais de primeira ordem de múltiplas maneiras, o que possibilita a capacidade de reflexão.

Mas, o que ocorre quando somos conscientes dessas representações? A consciência das representações parece ser simplesmente um dado. Ela não admite variações e manipulações intencionais, ao passo que a atenção, memória e imaginação as admitem. Quando nos voltamos introspectivamente podemos distinguir, intencionalmente, um número bastante rico de propriedades dos estados mentais. Como veículo da intencionalidade, a consciência é uma poderosa ferramenta cognitiva. Contudo, ao colocá-la como objeto não

conseguimos apreender-lhe as propriedades. A pobreza de nossa descrição é insanável. Para McGinn isto indica que a consciência possui uma estrutura oculta a si mesma; por um misterioso mecanismo, ela não é capaz de 'dobrar sobre si mesma', por assim dizer.

É também um fato, argumenta ele, que não podemos ter acesso a todas as experiências subjetivas existentes na natureza por meio de autopercepção. Os limites da nossa mente impõem constrangimentos aos tipos de experiência a que podemos ter acesso. Por exemplo: um morcego tem experiências fenomênicas cognitivamente fechadas para nós (a orientação por faixas de som que não conseguimos nem ouvir o revela).

Assim, se compreendêssemos T , teríamos uma propriedade explanatória geral capaz de elucidar como experiências fenomênicas ligam-se às suas bases físicas. Compreenderíamos também como a experiência do morcego liga-se ao seu cérebro. Consequentemente, ela seria capaz de explicar uma experiência que não possuímos e, portanto, revelar algo da natureza da experiência do morcego. Como esta consequência é impossível, logo, o antecedente é falso. É a tese 2 do argumento de McGinn (T2MA)

Analisando o outro lado do dilema, McGinn sugere que alguns proponentes de soluções ao problema mente-corpo poderiam apelar para a neurociência. Ao invés de investigar a mente introspectivamente, ou buscar uma teoria que desse acesso a P , eles poderiam voltar-se ao neurocientífico.

Novos estudos tem sido realizados acerca do tema. Não poderia P ser compreendido graças a maior extensão de nosso entendimento a respeito do funcionamento do cérebro? McGinn nega, partindo da idéia, muito razoável, de que propriedade da consciência não são *diretamente* observáveis na experimentação científica, não importa quão sofisticados sejam os métodos empregados.

Com efeito, que tipo de conhecimento se pode obter através de um estudo científico do cérebro? Observando o cérebro a olho nu vemos um órgão de tamanho médio com uma aparência fibrilada e cheio de circunvoluções. Podemos discriminar os componentes do encéfalo, os vários órgãos do sistema nervoso central, sua posição, partes, tirar amostras de tecido, etc. Descendo na

escala de tamanho, a aparência do cérebro é a de um conjunto enormemente numeroso de filamentos unidos uns aos outros com bulbos e por meio dos quais impulsos eletromagnéticos são transmitidos. Um dos modos de auferir informações da relação entre consciência e cérebro é através de seu escaneamento. Contudo, o que este escaneamento é capaz de fornecer é apenas um correlato de áreas mais estimuladas.

Como para ele a percepção tem um papel essencial na formulação de uma teoria sobre a mente, então sendo impossível perceber a propriedade da consciência em si mesma, logo é impossível formular a teoria M. Em outras palavras, a consciência é "noumenal a respeito da percepção do cérebro". (MCGINN, 1997,)

O próximo argumento tem a seguinte estrutura: McGinn mostrará que a propriedade P que causa a consciência é cognitivamente fechada e que não existe meio de inferir do que é percebido, P . Caso ele obtenha sucesso, então ficaria provado que P , não sendo alcançado por percepção ou inferência, não pode ser alcançado *de nenhum modo*. Chamemos de T3MA.

Argumenta-se, de início, que não existe nenhuma propriedade P que possa ser percebida como sendo causadora da consciência; e lança um desafio a que possamos apontar tal propriedade. Isto acontece porque as nossas teorias são apoiadas de maneira fundamental em percepções sensíveis, seja aferindo-as diretamente por meio da experimentação, seja postulando teoricamente uma propriedade que, embora não observável, é uma extensão analógica das primeiras.

Este argumento repousa em uma concepção da relação entre mente e mundo moldada de acordo com nossa história evolutiva. Novamente, McGinn traça uma analogia (MCGINN, 1999) com a imagem de uma criança tentando aprender, sem sucesso, a linguagem marciana. Esta insolúvel dificuldade provém do fato de que a gramática implícita estruturalmente na mente da criança não é compatível com a da linguagem marciana. O mesmo, diz o autor, talvez aconteça entre a gramática das ciências da natureza e aquela de uma aspirada "ciência da consciência".

Um dos traços mais básicos do pensamento é a combinação. Esta capacidade de combinação fica patente quando pensamos que as entidades complexas do mundo físico são compostas de entidades mais simples; que as leis mais específicas se subsumem às mais gerais; que partículas de linguagem como substantivos, adjetivos e conectivos juntam-se para formar sentenças. Este caráter combinatório do pensamento humano advém de duas fontes centrais: da linguagem e da percepção espacial. A teoria fisicalista do mundo nos fornece a mais clara mostra de um pensamento combinatório: entidades básicas como subpartículas atômicas compõem o átomo, que se combina até as moléculas e assim por diante chegando a organismos vivos ou grandes estruturas inorgânicas como estrelas, planetas e nebulosas.

O que está em jogo é a concepção de uma teoria científica: teorias científicas são baseadas em a) percepção de propriedades instanciadas espacialmente (a extensão macro e micro de um telescópio ou um microscópio apenas ampliariam o espectro do observável) ou b) em propriedades postuladas teoricamente que são *análogas* às primeiras. Tanto as primeiras quanto as segundas combinam-se numa sintaxe do mundo físico, do qual o cérebro é uma parte.

Mas o conhecimento da introspectivo da consciência, em seu caráter fenomênico, não converge com esta índole combinatória. A consciência não é o resultado da combinação de neurônios, ainda que a combinação de neurônios *instancie* a consciência. Uma percepção do vermelho da maçã não *consiste* na combinação de certos grupos de neurônios ativos. É possível deduzir analiticamente dimensões da experiência do vermelho como o matiz, o espaço, luminosidade, etc.,. De maneira idêntica, é possível decompor as dimensões complexas de uma percepção introspectiva refinada em elementos mais simples. No entanto, não se pode fazer isso da experiência para os seus componentes neurais. O que seria, afinal, a decomposição da experiência do amarelo na experiência do neurônio? Logo, a análise combinatória encontra limites quando passa do âmbito interno até seus hipotéticos constituintes materiais.

A sintaxe combinatória que serve tão bem para ciência física, gramática, química, biologia, e assim por diante, não funciona quando o desafio é compreender a consciência. Uma propriedade explanatória P não poderia ser a)

observada b) um análogo de observação. A comparação que McGinn invoca cita a formação do conceito de moléculas (MCGINN, Colin, 1997): formamos o conceito de moléculas reduzindo cada vez mais escala das medições até uma escala em que as entidades não podem ser verificadas diretamente. Entretanto, o caso da consciência é muito mais radical, uma vez que a fenomenologia da experiência não oferece nenhuma analogia escalar.

Por esta razão, mesmo que tivéssemos uma intuição de *P*, ainda assim haveria a mesma sensação de que algo permanece misterioso na relação cérebro-consciência. A incapacidade da mente de formar conceito de *P* é, porém, relativa. McGinn especula que uma outra mente, possuidora de conhecimento matemático do cérebro e da consciência talvez não estivesse sujeita às mesmas limitações que a da espécie humana. Isto porque o conhecimento matemático, ele afirma, não parece derivar de extrapolações da percepção sensível ou da faculdade introspectiva. Contudo, não é o conhecimento que podemos ter da consciência ou do cérebro. Chamemos essa de tese 3 do argumento (T3MA).

1.1.2 Problemas nas Teses Misterianistas

Temos, porém, algumas razões para considerar as duas últimas teses problemáticas. Essas razões devem nos levar à seguinte conclusão, caso estejamos corretos: o problema mente-corpo *pode*, em princípio, ser elucidado.

Em relação a T1MA, concordamos. A mera introspecção parece não ser capaz de fornecer nenhum tipo de conhecimento da relação entre a atividade de si mesma ou os seus conteúdos fenomênicos e o cérebro que a instancia. Ainda mais radicalmente: a hipótese cética de que somos um cérebro numa cuba não nos parece poder ser resolvida por meio da simples introspecção. Não há nenhuma evidência introspectiva decisiva em favor da saída do dilema cético, isto é, uma evidência tal que diminua a plausibilidade de que o cérebro simplesmente *não exista*. Se assim o é, então também não se pode obter conhecimento da relação consciência-cérebro apenas por evidência interna.

Por outro lado, em conexão com o caráter fenomênico da consciência, McGinn desenvolve um arrazoado sobre a sua simplicidade. Diz ele que a consciência não é o pináculo evolutivo e não parece ser mais complexa, por exemplo, que a linguagem (MCGINN, 1999, p.62). Parece-nos uma observação inteiramente aceitável. De fato, a consciência é compartilhada por animais de sistemas neuronais muito primitivos, ao contrário da linguagem articulada, que emerge, em sua estrutura sintática formal, apenas nos seres humanos. Ademais, podemos analisar a experiência fenomênica e perceber que, em termos de complexidade, ela não supera a própria intrincada mecânica neural que é, porém, acessível a investigação empírica e cujos detalhes são, gradativamente, mapeados pela ciência cognitiva. Ora, se a consciência nada tem de particularmente complexo, porque ela seria tão inabarcável?

Essa pergunta está ligada a este caráter fenomênico, assim como a questão do traço combinatorial do pensamento. McGinn utiliza essas duas intuições como pano de fundo para uma explicação alternativa: ele explica a dificuldade de se compreender a consciência como derivada de uma limitação conceitual de nossa mente. Essa é a tese principal em torno da qual gravita todas as subteses. Como teríamos uma dificuldade intrínseca de entender a consciência, oriunda de uma limitação primária na nossa mente, mesmo a sua simplicidade não facilita em nada a compreensão.

Acontece que essa explicação não parece particularmente elucidativa. Há no texto um salto conceitual curioso a conferir certa sensação de "wishful thinking" à conclusão de McGinn. Isto pode ser devidamente notado quando contrastamos esta explicação com alguns fatos tais como a existência de outros fenômenos desconcertantemente simples mas de explicação muito árdua. Se a dificuldade de explicar o fenômeno simples da consciência é um indício de que possuímos uma limitação cognitiva quanto a este objeto, então seria de esperar que dificuldade semelhante em explicar outros fenômenos simples fosse indicativo desta mesma limitação. Mas não é o caso.

Um desses fenômenos simples e de árdua explicação são certos problemas matemáticos fortemente contraintuitivos. O chamado problema de Monty Hall, famoso na TV desde a década de 70, ilustra um deles. Ele consiste, basicamente, no seguinte:

Um apresentador está com um convidado diante de três portas. As três portas estão inicialmente fechadas. Há um prêmio em uma delas, e dois monstros nas restantes.

- 1) O convidado é chamado a escolher uma das portas fechadas
- 2) O apresentador abre uma dessas portas, onde sabidamente está um monstro.
- 3) Sobram exatamente duas portas fechadas: a que foi escolhida pelo convidado e a outra.

A pergunta é: faz diferença permanecer na mesma porta ou é melhor mudar de porta? Normalmente, a resposta intuitiva é que tanto faz. Há somente duas portas fechadas, portanto, a probabilidade de o tesouro estar em uma delas é de $1/2$. Contudo, essa resposta (fortemente intuitiva) está completamente equivocada. Houve um famoso episódio envolvendo, de um lado, a genial Marilyn vos Savant, e, de outro, inúmeros matemáticos, inclusive importantes figuras na área da probabilidade. Ela ficou famosa por escrever, em sua coluna na década de 90, a resposta adequada a essa questão tendo recebido, por conta disso, numerosas críticas de matemáticos profissionais que simplesmente eram incapazes de compreender o problema por conta de seu forte componente contraintuitivo.

O problema de Monty Hall ficou famoso como exemplo de inadequação cognitiva diante de certos tipos de questões matemáticas. Mas, o que isso tem a ver com a argumentação de McGinn? O que chama a atenção, neste caso, é que se trata de um exemplo análogo. A formulação do problema é bastante simples, mas desafiadora aos nossos poderes intuitivos. Do mesmo modo, o fenômeno da consciência parece desafiar nossas intuições categoriais. Acontece que no primeiro caso a solução foi efetivamente descoberta. Mas é de se imaginar se, caso o problema não tivesse sido resolvido até hoje, que ele seria utilizado para mostrar que as capacidades cognitivas estavam intrinsecamente aquém de resolvê-lo. O que a argumentação de McGinn sugere é precisamente este último salto sem nada a ampará-lo exceto a (suposta) ausência de uma explicação satisfatória para a consciência.

Uma objeção que se pode fazer a isso é dizer que, tendo o problema sido resolvido, para a minha resposta se sustentar deveria ser necessário mostrar um sinal de resolução da questão mente-corpo. Para se resguardar dessa réplica, cumpre lembrar de um outro caso famoso na história da matemática, um dos problemas clássicos da geometria grega: o da trissecção do ângulo em partes iguais.

A formulação deste problema é extremamente simples. Seja um ângulo dado, divida o ângulo em três ângulos iguais. Mais de dois mil anos se passou para termos, no século XIX, a demonstração de que esse problema não poderia ser resolvido por meio de régua não graduada e compasso. Somente a ligação entre álgebra e geometria permitiu que se intuísse o modo adequado de resolver este problema. Evidentemente, não se tratava de uma mera ausência de instrumentos técnicos necessários, mas da muito mais crucial ausência de categorias conceituais adequadas para se solucionar o problema.

Este exemplo mostra que as categorias conceituais podem levar séculos e mesmo milênios para serem desenvolvidas até a maturidade suficiente onde certas questões possam ser desvendadas. Não existe nada que indique que a 'simplicidade' do fenômeno consciente é mais intrinsecamente elusiva do que a simplicidade desses problemas geométricos clássicos. Afinal, a filosofia da mente, em larga medida, ocupa-se desta questão. É possível que após séculos ou milênios, alguma individualidade genial atine com a categoria conceitual adequada para entender o fenômeno da consciência. Como a argumentação de McGinn tenta precluir essa possibilidade a priori, deve-se rejeitá-la, tomando como base questões análogas sob aspectos relevantes da simplicidade e |contraintuividade.

Já T2MA possui um problema lógico. É verdade que somos limitados a ter introspecção das experiências fenomênicas que possuímos; experiências de espécies distintas estão fora de nosso âmbito, por definição. A rigor, a experiência de nossos irmãos de espécie também. O constrangimento conceitual é, de certo modo, tautológico: não podemos ter introspecção daquilo que não é *nossa* introspecção.

O problema da conclusão de McGinn não é esta premissa óbvia, mas sim que ela salta por uma premissa que nos parece pouco clara. Qual a razão

de se imaginar que a posse de uma propriedade explanatória geral P desse meios de compreender introspectivamente o conteúdo das experiências cujo nexó físico seria explicado? Podemos perfeitamente imaginar uma explicação onde P mostrasse, em geral e abstratamente, como as experiências fenomênicas ligam-se a suas bases causais (tanto o *locus* físico dessa ligação quanto sua articulação neural e descrição do processo de emergência) sem que houvesse compartilhamento de experiências. Aliás, as tentativas de explicar a mente de maneira não reducionista são assim. Logo, a premissa não se sustenta por si mesma. Portanto, a conclusão não é cogente.

Finalmente, T3MA padece de dois problemas substantivos. Não parece verdade que *todo* conceito da ciência seja um análogo perceptivo. Existem paradoxos na física que não possuem nenhum análogo perceptivo, mas, ao contrário, contradizem a percepção. A utilização de geometrias de n dimensões na construção de diversas teorias físicas mostra que a percepção dessas dimensões, embora completamente impossível introspectivamente (e também sua observação) não impede que elas tenham uma função crucial em teorias que explicam dados observacionais.

McGinn poderia argumentar que, neste caso, os espaços hiperdimensionais só fazem ampliar imaginativamente a nossa percepção do espaço. Contudo, não é assim. A analogia com as moléculas é frágil porque os espaços de n dimensões são *inconcebíveis* exceto na condição de construção matemática, ao passo que as moléculas são apenas não observáveis.

A analogia forte vigora entre espaços de n dimensões e as capacidades auditivas dos morcegos. Do mesmo modo que é possível imaginar seres que possam perceber mais dimensões do espaço, também podemos imaginar (neste caso, existem realmente) seres que percebem frequências sonoras inacessíveis. Se alguma espécie animal desse indícios de que sua percepção do espaço atinge dimensões *extra*, poderíamos utilizar a geometria de n dimensões para descrevê-la estruturalmente, ainda que não tenhamos jamais nenhuma intuição do que elas percebem.

O outro problema é que McGinn descarta muito rapidamente a hipótese de algum tipo de conhecimento matemático da relação entre cérebro e consciência. A falha reside em considerar necessário que tenhamos

conhecimento matemático de proposições introspectivas ou da constituição do cérebro. Como nossa mente não possui (talvez a mente de Deus os conhecesse assim, especula McGinn), *então* essa teoria matemática é impossível. Não nos parece claro a relação de implicação entre essas proposições. Uma explicação que manejasse o instrumental matemático e computacional para mostrar como o correlato neural dá origem a consciência simplesmente *não precisaria* explicar a constituição do cérebro matematicamente (por completo) ou dar um tratamento matemático às próprias experiências fenomênicas.

A formulação de David Chalmers assume alguns pressupostos comuns com McGinn. Ele propõe uma teoria dualista para explicar a experiência fenomênica da consciência. Para sustentá-la, de um lado, desenvolve ampla argumentação positiva; de outro, enfileira uma série de contra-argumentos às teorias funcionalistas e reducionistas, particularmente contra as teorias neurobiológica de Crick e Koch, do espaço global de Baars, e das múltiplas camadas de Dennett (cf. TEIXEIRA, 1997), que buscam explicar a consciência em termos dos mecanismos que a permitem e são por ela implicados. Ao invés de buscar umnexo causal para o surgimento da consciência, ele prefere limitar-se a apontar um correlato neural. Deste modo, escapa da dificuldade corretamente assinalada por McGinn segundo a qual as explicações físicas do cérebro compõem um circuito causalmente fechado.

O núcleo da proposta de Chalmers (CHALMERS, 1996, 2010) é a afirmação de um tipo de sobreveniência da consciência em relação ao cérebro. Para ele, as explicações dos processos mentais dividem-se em dois gêneros: explicações sobre os processos mentais em sentido psicológico e explicações sobre a experiência consciente. Aos problemas teóricos relativos aos processos mentais em sentido psicológico ele denomina *easy problems*, e dá como exemplo questões a respeito do mecanismo de memória, aprendizado de uma linguagem, etc. Ao segundo, chama-o *hard problem* e menciona a questão a respeito da natureza dos *qualia*, o substrato perceptivo de acesso imediato que temos quando temos uma experiência: a visão de uma cor, o tato de uma superfície, etc. Este *hard problem*, porém, não é essencialmente "misterioso".

A dificuldade dos problemas é graduada conforme o tipo de questão envolvida. No caso dos *easy problems*, por complexo que seja sua resolução, ela se dá no marco da investigação usual das ciências cognitivas e, filosoficamente, concerne a processos mentais explicáveis pelos instrumentos conceituais de posse dos funcionalistas. Não há nenhum *desafio metafísico* para a explicação usual dos processos mentais. Explica-se um processo mental de acordo com o papel que eles desempenham na produção de um comportamento. Já o *hard problem* traz este desafio, uma vez que ao explicar a experiência como o resultado de um estímulo do ambiente e do processamento do estímulo pelo cérebro contorna-se a questão fundamental. A questão fundamental é: porque há um correlato fenomênico do processo psicológico, e o que é este correlato? Observe-se que o processo de processamento psicológico é um processo não consciente. A percepção é consciente, e é imediata como um *dado*. Explicar o processo não é idêntico a explicar o dado imediato da consciência, logo, a explicação que apela exclusivamente a esclarecer o processo psicológico é insuficiente.

1.2. SENTIDO PSICOLÓGICO E SENTIDO FENOMÊNICO

Em *The Conscious Mind*, (CHALMERS, 1996, p.11-16) estabelece uma distinção crucial para o resto de seu trabalho. Essa seção, por seu turno, é um subtópico de um capítulo intitulado "The Two Concepts of Mind". Esse título é, decerto, uma alusão ao clássico de Gilbert Ryle, *A Concept of Mind* de 1949, que formula a posição behaviorista no debate contemporâneo da filosofia da mente.

É trivial constatar que um indivíduo consciente não realiza todas as atividades mentais "no escuro", à maneira de um autômato. Atribuímos aos indivíduos um âmbito interno subjetivo, no qual experimentam a si mesmos como entes conscientes, e experimentam alguns conteúdos da sua consciência. Exemplos de conteúdos experimentados são as sensações e os pensamentos. Se eu acerto um martelo no meu dedo, e está tudo em ordem com meu sistema

nervoso, eu experimento a sensação "dor". Se eu olho pela janela, vejo a chuva e penso "chove lá fora", estou consciente de que penso "chove lá fora".

Isto significa que algo se passa em meu âmbito subjetivo; sou consciente de algo, portanto, não executo todas as atividades mentais "no escuro". Há também um tipo característico de acesso a esse âmbito subjetivo, que nos permite aceitar a crença de que temos um âmbito interno pelo mero autoexame reflexivo, sem precisar recorrer a provas suplementares. Analogamente, atribuímo-lo a outros indivíduos, cujo comportamento assemelha-se suficientemente ao que exibimos quando experimentamos os conteúdos desse âmbito subjetivo para permitir uma atribuição confiável. Muito embora o desafio cético tenha fornecido razões relevantes para se duvidar de que essa atribuição seja realmente confiável, temos, contudo, uma forte intuição a favor dela. E julgamos a crença na existência de nosso próprio âmbito subjetivo muito mais sólida, em virtude do acesso direto que temos a ele. As opiniões filosóficas a respeito da segurança epistêmica do referido acesso variam significativamente, é verdade, mas, geralmente, não se desautoriza a confiança de que, ao menos, *temos* um âmbito subjetivo.

Por outro lado, as ciências cognitivas lidam com estados condicionantes de comportamento,

A experiência consciente não é tudo o que diz respeito à mente. Para ver isto, observe que, embora a ciência cognitiva moderna não tenha tido quase nada a dizer sobre a consciência, ela tem muito a dizer acerca da mente em geral. Os aspectos da mente com os quais ela se importa são diferentes. A ciência cognitiva lida largamente com a explicação do comportamento e, enquanto preocupada com a mente como um todo, interpreta-a como a base interna do comportamento, sendo os estados mentais relevantes à causa e à explicação do comportamento (CHALMERS, 1996, p. 11)

Ele diferencia de modo nítido o sentido fenomênico do sentido psicológico da mente, e o faz na seguinte chave: o conceito fenomênico da mente concerne à experiência em si mesma, não a tomando como *fator causal* na produção de um comportamento. Em oposição a este, o sentido psicológico é a base explanatória ou causal para o comportamento.

Chalmers faz questão de esclarecer que essa é uma distinção conceitual. A ele lhe parece perfeitamente possível, de início, que o aspecto ou

sentido psicológico da mente tenha a mesma referência que o sentido ou aspecto fenomênico. Contudo, a distinção conceitual existe ainda assim.

Um exemplo, extraído de *The Character of Consciousness*, pode ajudar a esclarecer a significação de conceito fenomênico. (CHALMERS, 2010, p. 254). Suponha-se um sujeito que vê uma maçã vermelha e experimenta visualmente sua cor. Esta experiência instancia uma qualidade fenomênica V, a vermelhidão da maçã. Há, no caso, duas qualidades: uma é a vermelhidão da experiência, a outra a superfície do objeto exterior cuja experiência foi efetuada. A vermelhidão da experiência é um conteúdo da experiência consciente, algo que a mente *sente*, não *realiza*.

Explorando o exemplo, pode-se notar que as condições de possibilidade da experiência fenomênica abrangem várias tarefas que a mente executa no processo de experimentar: o foco da atenção no objeto percebido, permitindo ao sujeito do exemplo “notar” a maçã a partir do entorno físico; a interpretação do estímulo recebido através do olho pelo cérebro; o acesso a este estado interno; a retenção do dado na memória. São estas atividades que compõem o correlato psicológico da experiência fenomênica. Contudo, elas não esgotam o ato perceptivo. O resultado do processo mental é um tipo de experiência. Essa experiência não está determinada pelo processo; aparece como um dado dependente, porém irreduzível do processo.

Esses conteúdos da experiência são inventariados de forma não exaustiva: cores, formas, cheiros, a audição de uma música, pensamentos e o senso do "Eu" são alguns dos exemplos mencionados. A noção de experiência envolve a noção de um sujeito que *tem* um estado, que *sente* algo.

Assim, é traçada a história da abordagem do problema da consciência detendo-se em duas opções explicativas paradigmáticas, opostas. Por um lado, a explicação cartesiana amplia o seu âmbito a ponto de considerar que todo estado mental é estado consciente. Por outro lado, a explicação da mente em termos puramente psicológicos, no sentido especial do termo, tenta explicar a experiência consciente de forma igual ao modo como explica os demais processos mentais, sem considerar aquilo a diferencia dos demais processos.

A despeito da breve menção a Gilbert Ryle, ao qual ele credita a originalidade de haver primeiro rompido com a concepção cartesiana, Chalmers

prefere enfrentar a abordagem funcionalista por considerá-la mais refinada. A vantagem explicativa do funcionalismo sobre Ryle concerne a capacidade do primeiro em escapar de duas objeções. Segundo a teoria de Ryle, os estados mentais são estados comportamentais. Duas objeções (dentre muitas outras possíveis, que já foram aventadas na introdução) podem ser feitas a essa asserção. Primeiro, se os estados mentais são estados comportamentais, então eles não podem *causá-los*. Segundo, nenhuma disposição comportamental pode ser explicada em termos de um único estado mental. Se alguém acredita que chove, suas disposições comportamentais dependerão de se ele quer se molhar ou não. Do mesmo modo, ele querer se molhar ou não poderá depender de ele acreditar que ficará resfriado ou não. E assim por diante.

A explicação funcionalista elide estas dificuldades, na medida em que cada estado mental é definido de acordo com seu papel causal, com o estímulo comportamental que tende a provocar e com a relação que guarda com os demais estados. Com efeito, Armstrong (1968, p. 85) opera uma interessante distinção entre duas concepções de disposição. De um lado, existe a concepção fenomenológica de disposição, a qual ele atribui a Ryle e a outros behavioristas. Segundo esta concepção, um estado disposicional não possui necessariamente um estado categorial que lhe sirva de base, ou seja, possuir uma propriedade disposicional não é estar em um estado x , mas apresentar um comportamento envolvido na descrição da disposição quando satisfeitas as condições de sua manifestação. De outro lado, Armstrong argumenta em favor da concepção realista das disposições, segundo a qual elas exigem um estado categorial de base.

Para provar a veracidade de seu ponto, Armstrong fornece o seguinte exemplo: imagine-se uma borracha, que sob força F move-se uma polegada. Isto acontece todas as vezes que se aplica a força. Pode-se presumir, portanto, que no tempo $t1$ a borracha se moveria uma polegada, ainda que neste tempo específico nenhuma força lhe esteja sendo aplicada. Este exemplo traz uma série de dificuldades para quem tem a concepção fenomenológica. Se a borracha não está sob a força F , e não existe nenhum estado categorial de base, então em um tempo tn onde ela não esteja sendo objeto dessa força não podemos afirmar que ela se afastaria uma polegada.

O erro da abordagem funcionalista consiste em querer analisar todos os estados mentais desde esta plataforma conceitual. Segundo o autor, o sentido fenomênico não se presta a ser capturado por ela. Como o que importa para a fenomenalidade do estado considerado é o modo como ele é *sentido*, e não a sua função, a abordagem de Lewis e Armstrong revela-se incongruente com a fenomenalidade.

Isto o leva a postular uma nova dimensão explicativa para a consciência. Do mesmo modo que a física possui leis e conceitos básicos – massa, spin, carga, e assim por diante – a experiência fenomênica também possuiria princípios básicos próprios. Evidentemente, esta dualidade no conjunto de princípios não é tal que ou exija a refutação das explicações oriundas das ciências cognitivas, ou a consequente adesão a uma metafísica não naturalista da mente.

Ele aposta na validade do modelo epistemológico das ciências cognitivas como também advoga claramente uma cosmovisão naturalista. Neste sentido, a sua obra é afim às cogitações de McGinn. Então, se o modo naturalista é válido, a relação de dependência entre o fator singular da consciência e a estrutura material que a suporta, explicada por este modelo, mostra-se particularmente crucial para a solidez da sua teoria.

Para Chalmers (1996), a relação entre a consciência e o cérebro é desse tipo, o que significa que se pode afirmar que não há aspecto qualitativo do estado mental consciente sem que haja correspondência estrita no sistema nervoso mas, que sua contrária é concebível. Com efeito, uma série de experimentos mentais tem seu tratamento unificado a partir do argumento da conceitabilidade e constitui um dos principais argumentos a favor do dualismo de Chalmers. No núcleo do argumento, contudo, está a noção de sobreveniência, que será tratada no capítulo seguinte.

2 SOBREVENIÊNCIA E O PROBLEMA DA CAUSAÇÃO

2.1 CONCEITOS DE SOBREVENIÊNCIA

Atribui-se a G.E. Moore a primeira elaboração filosófica relevante, em termos contemporâneos, da idéia de sobreveniência. A R. M. Hare credita-se a utilização pioneira do vocábulo correspondente (HARE, 1952). Apesar das diferenças nas formulações, eles compartilham a preocupação em responder a questões similares, relativas à ontologia dos valores. Recentemente, (HORGAN, 1993, p. 555-565) defendeu que os emergentistas britânicos, destacando-se, entre eles, Morgan, empregavam a noção com a acepção vigente no contexto do debate evolucionista, indicando precedente histórico mais remoto. Em todo caso, qualquer que seja a origem da expressão, dois exemplos de *The Language of Morals* (HARE, 1952, p.79-85) auxiliam a extração do núcleo semântico que vigora no cenário atual.

Suponha estar o quadro *Vista de Delft*, do holandês Veermer, exposto em um museu. Um crítico julgará, com acerto, tratar-se de excelente pintura. Imagine-se uma cópia idêntica, feita pelo mesmo autor na mesma época em que pintara a homônima. Vendo-a, nosso juiz hipotético diz "Vista de Delft é uma excelente obra. Não difere de sua cópia em nada exceto por seu valor superior". Facilmente nos ocorreria a réplica: "Mas, se elas são iguais em tudo, o que faz uma ser excelente e a outra inferior?" Ao que o hipotético avaliador argumentaria: "Não disse que diferiam em algo além de uma ser excelente e a outra, inferior. Do mesmo modo, eu poderia dizer que uma foi assinada e a outra, não, e que essa era a única diferença. Então, uma poderia ser excelente e a outra, inferior, e não diferirem em mais 'nada'".

O segundo exemplo converge com o anteriormente exposto: um novo juiz hipotético diz "São Francisco é um bom 'homem'". Suponha outro sujeito, S₂, cujo comportamento fosse idêntico ao de São Francisco, nas mesmas circunstâncias; neste contexto, avaliando-o externamente e internamente igual, o avaliador afirma "S₂ comporta-se do mesmo modo que São Francisco em todas as circunstâncias relevantes. Contudo, S₂ não é 'bom'".

Julgamentos como os descritos parecem inaceitáveis devido ao fato de concebermos, intuitivamente, a seguinte condição: para que o valor estético ou moral de dois objetos, tomados sob mesmo aspecto, possam ser distintos, é necessário haver uma diferença *extra* entre eles, ou em suas circunstâncias, relevante para a determinação de tais valores. Deste modo, é correto dizer que a distinção de valor *depende* de uma disparidade adicional entre os objetos ou entre suas circunstâncias, ainda que, segundo Hare, não exista uma conjunção de proposições descritivas que implique logicamente uma proposição valorativa. Mas, não é necessário que esse conjunto exista para que haja uma relação de sobreveniência.

A noção em pauta, a partir de definições mais específicas presentes em autores como Moore ou Hare, pode ser generalizada formalmente e aplicada a um conjunto bastante variado de elementos. Na filosofia da mente, em especial, os trabalhos de D. Davidson (1970) e J. Kim (1984), entre outros, revelaram-na profícua, ao mostrarem que as relações entre os estados mentais e os estados cerebrais podiam ser explicadas em termos de relações de sobreveniência dos primeiros sobre os últimos. J.Kim:

(KIM, 1984, p. 64) dá algumas definições clássicas. Sobreveniência fraca é definida do seguinte modo,

"A fracamente sobrevém a B se e somente se necessariamente para quaisquer x e y se x e y compartilharem todas as propriedades em B então x e y compartilham todas as propriedades em A - isto é, indiscernibilidade em relação a B implica indiscernibilidade em relação a A.

Kim vale-se do caso de São Francisco, de Hare, a fim de ilustrar sua noção. Procederemos do mesmo modo com o nosso exemplo do quadro de Veemer. Suponha-se que o quadro possua um conjunto base de propriedades-B, sendo esse formado por três propriedades relativas à qualidade da execução da pintura: T, técnica, C, criativa, E, expressiva. Já o conjunto sobreveniente de propriedades-A contém o elemento Ex, excelente. Consideremos que dois conjuntos-B tenham as mesmas propriedades. Assim sendo, o conjunto sobreveniente para cada um deles será o mesmo, contendo Ex ou ~Ex. Em contrapartida, distinções em dois conjuntos A implicam disparidades nos

respectivos conjuntos-B, aos quais os primeiros sobrevivem. Destarte, se dois quadros possuírem um mesmo conjunto maximal de propriedades-B (T, ~C e E), ou (T, C e ~E), ou (T, ~C, e ~E) eles serão indiscerníveis quanto ao conjunto-A.

Tal relação ainda é demasiado fraca para que se possa dizer que as propriedades de base garantem ou determinam as sobrevenientes. Dois quadros que possuam o mesmo conjunto maximal, e.g. (T, ~C e E) têm o mesmo conjunto-A. Pode-se, entretanto, conceber uma situação em que esse conjunto maximal não determine o mesmo conjunto-A. Imagine-se um mundo possível onde essas mesmas características não impliquem o valor de um quadro, não havendo nenhuma relação de cogência entre B e A. Apesar disso, a relação de sobreveniência fraca poderá ser preservada, denotando que ela não é cogente.

Não obstante, uma teoria fisicalista, segundo Kim, deve ter pretensões de mostrar que as propriedades mentais são determinadas por propriedades físicas, e não apenas que da identidade de conjuntos de propriedades-B se segue a identidade de conjuntos de propriedades-A¹.

Deste modo, (KIM, 1984, p.65) prossegue fornecendo a definição de sobreveniência forte,

A fortemente sobrevém a B apenas no caso de, necessariamente, para cada x e cada propriedade F em A, se x tem F então há uma propriedade G em B tal que x tem G e, necessariamente, se qualquer y tem G, ele tem F.²

A diferença entre os dois tipos de sobreveniência pode ser ilustrada utilizando o mesmo exemplo anterior do quadro de Veermer. No caso da sobreveniência fraca, dois conjuntos-B, formados pelas mesmas propriedades de técnica, criatividade e expressividade implicam dois conjuntos-A idênticos de excelência. Entretanto, é possível que um quadro y, possuindo as mesmas propriedades em

¹ Portanto, para Kim o monismo anômalo seria uma tese insuficiente, uma vez que, para salvaguardar a ausência de leis psicofísicas estritas, ela esposaria uma noção fraca de sobreveniência. A analogia de Davidson entre a relação do físico com o mental e a dos aspectos sintático e semântico de sentenças envolveria uma noção de tipo fraco; embora o aspecto semântico dependa da sintaxe, ele não é por ela garantida para todos os mundos possíveis. Em um mundo onde significados diversos fossem atribuídos aos mesmos elementos sintáticos, o conteúdo proposicional poderia ser diferente daquele no mundo original. Não há, pois, um liame de necessidade entre sintaxe e semântica.

² Essa definição é a revista na parte posterior do livro.

B de um quadro x; afinal se temos um x, tal que x seja um quadro, e este quadro possua a propriedade C e T (propriedades de tipo G), em B e Ex (propriedade de tipo F) em A, apenas pela sobreveniência forte outro quadro, y, necessariamente se tiver criatividade e técnica terá excelência.

É de se notar que a relação de sobreveniência caracteriza-se por ser assimétrica. Como salienta Kim, embora a tese segundo a qual o psicólogo sobrevém ao físico seja plausível, a tese contrária - o físico sobrevém ao psicológico - é trivialmente falsa. Por seu turno, relações de sobreveniência não implicam relações de dependência ou garantia em sentido forte, isto é: que as propriedades do conjunto-A sejam tais que o são *em virtude* das propriedades da base. A sobreveniência forte apenas indica que, uma vez dadas as propriedades do conjunto base, seguem-se as propriedades do conjunto superior; todavia, se elas seguem em razão daquelas não é uma conclusão que se possa extrair imediatamente. A conclusão é obtida mediante uma solução do autor, particular, e que não interessa diretamente a nossa exposição³.

Interessam-nos, todavia, o conceito de sobreveniência global, uma vez que ela reenquadra o tópico em termos de mundos possíveis, e não somente de propriedades instanciadas em particulares. Isso ocorre, posto sua definição

A globalmente sobrevém em B somente em caso de mundos que são indiscerníveis com respeito a B (B-indiscernível, em resumo) também serem A-indiscerníveis.(KIM, 1984 p. 68)

Essa definição pode ser lida, seja lançando mão da linguagem das propriedades e suas instanciações, seja falando em fatos. Assim, dois mundos indiscerníveis quanto ao conjunto de fatos B serão indiscerníveis no conjunto de fatos A: dois mundos iguais quanto a seus fatos físicos serão iguais quanto a seus fatos psicológicos, caso o mental sobrevenha globalmente ao físico. As duas abordagens são equivalentes porque um fato de tipo A pode ser descrito

³ Kim defende uma solução, segundo a qual propriedades superiores são determinadas, causalmente ou não, por propriedades base quando inseridas em um sistema holista. O exemplo da dependência psicofísica ilustra o caso: uma propriedade mental F, que sobrevém a uma propriedade física G (portanto, G em B, domínio físico sobre F em A, domínio mental) depende dessa propriedade pelo fato (putativo) do mental sobrevir ao físico, isto é, de um conjunto extenso de propriedades daquele domínio - ou o domínio completo - comportar-se de modo idêntico a propriedade individual considerada.

como um particular a , tal que possui certa propriedade, F em A e mundos como conjuntos maximais de fatos.

Algumas noções apresentadas (CHALMERS, 1996, pp. 32-46) guardam semelhanças notáveis com essa exposição clássica. Em primeiro lugar, Chalmers parte da constatação de serem os fatos físicos os mais básicos na estrutura do universo, dos quais os restantes dependem de diversas maneiras. Esta é uma presunção fisicalista, como assinalado por Kim, que também demarca o território onde o projeto filosófico dele se desenvolverá.

Tal dependência é qualificada. Poderíamos pensar, em um primeiro momento, em várias formas dela possíveis: ora uma dependência nomológica, em que leis que explicam fatos de baixo nível implicam leis que explicam fatos de alto nível; ora uma dependência causal, como a que se verifica no movimento de um objeto A e subsequente colisão com um B , provocando o movimento deste último; poderíamos ainda conceber um tipo de dependência constitutiva, como aquela que vigora na teoria da identidade (PLACE, 1954) (SMART, 1959). Esses tipos de dependência possuem alguma relação com a noção de sobreveniência, porém afirmam mais do que se pode extrair dela. Seguindo a mesma linha aberta por Kim, Davidson e outros Chalmers formula a definição básica em termos de propriedades,

Propriedades-B sobrevêm a propriedades-A se não há duas situações possíveis que sejam idênticas em relação às suas propriedades-A enquanto diferem em suas propriedades-B. (CHALMERS, 1996, pp. 33)

Após breve discussão a respeito da sobreveniência local e global, onde ele mostra que a primeira implica a segunda e não vice-versa, Chalmers faz a distinção mais importante do capítulo: diferencia os tipos natural e lógico de sobreveniência (CHALMERS, 1996 p. 34):

Um conjunto y é logicamente sobreveniente a um conjunto x de propriedades quando é logicamente impossível que, dado o conjunto x , não se dê y . Deste modo, se dois conjuntos de propriedades superiores (de tipo F , $F1$ e $F2$) são idênticos, então não podem ser distintos com respeito às propriedades de tipo G , $G1$ e $G2$, que formam seus respectivos conjuntos-base.

Ele intercambia livremente fatos por propriedades. Assim, se um conjunto $G1$ de fatos físicos são fixados em um mundo $w1$, e um outro conjunto $G2$ é fixado em outro mundo $w2$, então os conjuntos de fatos biológicos $F1$ e $F2$ são também fixados para aqueles mundos respectivos. Globalmente, se todos os fatos físicos de um mundo $w1$ estão dados, os fatos biológicos do mundo não podem variar; localmente, a relação de dependência estrita verifica-se igualmente: para cada conjunto possível G corresponde apenas um conjunto F . Ao passo que a sobreveniência lógica é constrangida por limites essencialmente conceituais, a sobreveniência natural é constrangida por limites naturais, isto é: fatos de um tipo sobrevém a fatos de outro tipo em virtude das leis de funcionamento de um mundo.

A lei dos gases fornece um caso deste último tipo. Ela estabelece que o valor da pressão de um gás pode ser dado por uma função a partir do valor da temperatura, do volume, e de uma constante k . A pressão dos gases sobrevém a temperatura e volume: variando os fatos da temperatura e volume, variam os fatos a respeito da pressão. A validade da equação que a descreve ' $pV = KT$ ' apoia-se na necessidade, empiricamente verificável, da relação descrita matematicamente na fórmula.

Entretanto, nada assegura que em outro mundo esta equação não tivesse uma variável diversa, e que a variação dos conjuntos de fatos fosse distinta. Outras equações físicas, previsivelmente, enquadram-se na mesma tessitura. Por exemplo: a fórmula que dá a atração gravitacional, consideradas as atuais condições do mundo, tem certos valores fixos. Imaginando um mundo possível n com diferente configuração física, a gravitação universal poderia ser descrita com outras constantes, ou mesmo não existir.

Observa-se a presença, em toda a explicação sobre a noção de sobreveniência, a qual segue de perto a formulação do texto original, da ideia de possibilidade. Por razões que desenvolverá mais profundamente em outros escritos, Chalmers intercambia os conceitos de possibilidade lógica e metafísica. Faz o mesmo com os de possibilidade e conceptibilidade cumprindo, grosso modo, a fórmula: o domínio do possível é o mesmo do que é concebível, e este último implica a possibilidade metafísica.

Neste sentido, os limites da sobreveniência lógica são dados por concepções que envolvem contradição lógica, já os limites da sobreveniência natural são dados por concepções que confrontam a configuração natural do mundo, mas são, logicamente, possíveis. O principal modo de demonstrar que a sobreveniência da consciência é natural é mostrar que é possível um mundo onde todos os fatos físicos do nosso mundo se deem e não exista nenhum tipo de fenomenalidade consciente.

2.2 ARGUMENTO BIDIMENSIONAL E PROPOSIÇÕES FENOMÊNICAS

Afirma Chalmers que o modo mais óbvio, embora não o único, de investigar a sobreveniência lógica da consciência é conceber um mundo fisicamente idêntico ao nosso, no qual, contudo, a consciência, em sentido fenomênico seja globalmente ausente. Essa estratégia coaduna-se com a sua compreensão de que a superveniência lógica não pode ser definida em termos de dedutibilidade em um sistema formal, mas em termos de uma metafísica de mundos possíveis. Caso vitoriosa, a estratégia resultará em mostrar que a consciência não sobrevém logicamente ao conjunto de fatos microfísicos.

Um dos modos de implementar essa estratégia é através do que se denomina argumento bidimensional contra o materialismo, também conhecido como *argument of conceivability* (argumento da conceitabilidade), cuja forma lógica inicial é a seguinte (CHALMERS, 2010, p 141):

1. $P \& \sim Q$ é concebível.
2. Se $P \& \sim Q$ é concebível, $P \& \sim Q$ é metafisicamente possível.
3. Se $P \& \sim Q$ é metafisicamente possível, o materialismo é falso.

-
4. Materialismo é falso.

Neste argumento, P é o conjunto de todas as verdades microfísicas do mundo, especificando o conjunto completo de entidades, aspectos e fatos microfísicos do mundo possível e, portanto, dando a sua configuração macrofísica, por sua vez determinada por aqueles fatos. Já Q é uma certa verdade fenomênica. Dois exemplos de verdades fenomênicas relevantes, respectivamente, para o argumento dos zumbis e o do espectro invertido: Q: alguém possui consciência Q': alguém instancia um tipo específico de propriedade fenomênica.

Se Q, P&~Q significa que neste mundo possível não há ninguém que possua consciência, isto é, trata-se de um *mundo de zumbis*. Se Q', P&~Q' significa que neste mundo possível há alguém que não instancia um tipo de propriedade fenomênica. Exemplo: alguém não instancia o tipo de propriedade fenomênica 'amarelo', e ao invés dela instancia 'vermelho', o que é relevante para o argumento do espectro invertido.

Chalmers nota que se Q for uma verdade negativa, tal como a de que não há alguém que instancia certa propriedade fenomênica *extra*, podemos assumir que há outros mundos possíveis onde P&Q completa todos os fatos do mundo, então P&~Q torna-se compatível com o materialismo. Para evitar essa consequência, conjuga-se T a premissa inicial, onde T significa "*that's all*", para indicar que se trata de um mundo *mínimo* que satisfaz P e, assim, salvaguardar verdades negativas. A inclusão de T significa que nosso P implica e só implica aquilo que existe em um mundo mínimo atual.

(CHALMERS, 2010, p. 109) desenvolve algumas considerações a respeito da noção de *conceivability* (conceptibilidade). S é concebível negativamente quando não se pode negar S *a priori*; positivamente, quando S é coerentemente concebível *a priori*. É concebível positivamente *prima facie* quando ao sujeito é impossível negar S por um raciocínio *a priori* inicial; concebível negativamente *prima facie* quando é impossível imaginar S coerentemente por um raciocínio *a priori inicial*, ao passo que é concebível negativa idealmente quando é impossível negar S por um raciocínio *a priori* válido, e concebível positiva idealmente quando conceber S coerentemente não pode ser negado por um raciocínio *a priori* válido. A combinação das quatro

possibilidades dá o quadro CP+ *prima facie*. CP+ ideal; CP- *prima facie*; CP - ideal.

Alguns argumentos cujo tratamento é unificado no supracitado. Aqui, apenas menciono sinteticamente:

1) O argumento do espectro invertido é extraído de uma nova situação contrafactual, sob alguns aspectos, semelhante à do experimento mental anterior. Neste, ao contrario do mundo de zumbis onde a consciência não existe, a experiência consciente existe, mas difere da experiência no mundo real.

A distinção é a seguinte: na situação contrafactual, há um mundo gêmeo onde as experiências são globalmente invertidas em relação ao original; localmente, isto implica que a experiência de um Ricardo gêmeo, fisicamente idêntico ao original, também é sistematicamente invertida. Se existe essa inversão, então os fatos fenomênicos podem variar independentemente dos físicos, o que mostra que a relação de sobreveniência é válida para aquele mundo, mas não atravessa todos os mundos possíveis. Trata-se, pois, de sobreveniência natural⁴.

2) O problema das outras mentes normalmente é apresentado como um desafio cético a justificação da crença segundo a qual existem outras mentes além da nossa. Ele surge em virtude de uma condição peculiar do acesso ao mental. Aparentemente, não é possível ter acesso aos estados fenomênicos, incluídos nestes os estados representacionais, senão por dois modos: evidência direta, no caso da nossa própria mente; e inferência. A evidência direta é obtida a partir de uma experiência interior imediata: quando penso que estou digitando no teclado nesse exato instante, sei imediatamente e com autoridade de primeira pessoa que o faço. A inferência pode se dar a partir dos nossos próprios estados mentais quando, por exemplo, recuperamos um conteúdo da memória por meio de associações. Mas, pode-se dar também para explicar o comportamento dos outros agentes humanos, como no caso de postularmos formas de consciência para terceiros.

⁴ Os três argumentos estão em (CHALMERS, 1996, pp. 93-103).

É fácil notar que, quando observamos outros seres supostamente dotados de uma mente temos acesso apenas a seu comportamento externo. Em um nível mais fundamental, podemos ter acesso ao funcionamento do seu cérebro, visualizando a complexa rede de interações neuronais que configuram os estados cerebrais de um indivíduo dotado de sistema nervoso. A partir desse dado, para alguns seres nós inferimos a existência de um âmbito de experiência consciente. Todavia, uma das consequências do argumento bidimensional é que, mesmo se pudéssemos, idealmente, fornecer uma descrição exaustiva das interações cerebrais, não teríamos com isso o acesso às próprias experiências desse indivíduo. Portanto, não temos como inferir a existência de outras consciências a partir do conhecimento objetivo dos estados cerebrais. A esse argumento denominamos argumento da assimetria epistêmica.

3) Imaginemos um mundo onde as neurociências tivessem alcançado tal grau de desenvolvimento de modo que toda a informação a respeito do processo físico no cérebro responsável pelo comportamento humano estivesse disponível à comunidade científica. Mary, uma notável neurocientista, possuidora de conhecimentos ilimitados sobre o processamento cerebral dos estímulos de cor, foi, no entanto, para seu infortúnio, criada em uma sala preta e branca. Nunca viu Mary outras cores exceto preto, branco e tons de cinza. Nesta circunstância, pergunta-se: ela pode obter saber o que é o 'vermelho' sem a experiência correspondente?

Uma vez que os conhecimentos dela acerca do funcionamento do cérebro são completos, então ela sabe o que acontece ao cérebro quando ele é submetido a certo comprimento de onda que produz nas demais pessoas a evidência sensível do 'vermelho'. De posse desses conhecimentos, ela obtém inferencialmente o conhecimento qualitativo do 'vermelho'? Aparentemente, a resposta é não. Se assim o é, então isso significa que as premissas do argumento bidimensional estão corretas e, portanto, a consciência não sobrevém logicamente ao sistema nervoso.

A validade do argumento bidimensional não está em jogo no presente trabalho. Existe extensa discussão a respeito na literatura⁵. O que importa observar é o traço que unifica a maior parte desses argumentos: a presença de dois “gaps”, o ontológico e o explicativo.

2.3 REDUACIONISMO E PRINCÍPIO DO FECHAMENTO CAUSAL

Há dois tipos básicos de redução das propriedades mentais às físicas. Por um lado, podem-se reduzi-las através da identificação das primeiras às últimas. Exemplo disso é a redução do calor ao movimento das moléculas. Trata-se de reducionismo ontológico. Por outro lado, podem-se reduzir as leis que regem determinado campo de fenômenos às leis de outro campo, por exemplo, quando se tenta reduzir leis biológicas a leis da química.

Teorias não reducionistas dividem-se, igualmente, em teorias que admitem o gap ontológico, e constituem assim uma sorte de dualismo de propriedades, ou teorias que negam o gap ontológico, mas admitem o explicativo. A teoria dualista de Chalmers exige os dois gaps. Não obstante, estranhamente, ela aceita um fechamento causal dos sistemas físicos. Cito-o,

Particularmente, uma teoria não-redutiva da experiência especificará princípios básicos que nos dirão como a experiência depende das características físicas do mundo. Estes princípios psicofísicos não interferirão em leis físicas, visto que, aparentemente, estas já constituem um sistema fechado. (CHALMERS, 2010, p.17)

Esse isolamento das leis psicofísicas recorda a ausência de leis psicofísicas estritas do monismo anômalo davidsoniano. Veremos que o fechamento causal conduz a algumas consequências contraintuitivas para teorias não reducionistas.

O princípio do fechamento causal estabelece que,

⁵ O capítulo VI de *The Character of Consciousness* é a melhor fonte a respeito. Ao final da exposição do argumento, Chalmers fornece extensa lista de objeções e responde a cada uma delas.

(PFC): Todo efeito físico possui uma causa física suficiente

Aceita-se geralmente que o princípio do fechamento causal é exclusivo da física. Em ciências que operam com propriedades de nível superior como, por exemplo, a biologia, o princípio não vigora. Como ilustra (PAPINEAU, 2008, p.54) quando o meteoro caiu no planeta e extinguiu os dinossauros, uma causa exclusivamente física produziu um efeito biológico. Do mesmo modo, causas físicas provocam efeitos mentais: uma aspirina alivia uma dor de cabeça forte, certos psicotrópicos geram alterações comportamentais.

Mas o inverso parece não ocorrer: a cadeia causal no reino da física aparenta estar fechada. Para cada efeito físico, parece haver um evento suficiente para explicar seu acontecimento, do qual é causa. Se uma infecção produz um aumento de temperatura, ou se os meus pensamentos movem os meus dedos no teclado, há explicações exclusivamente físicas para as duas situações. A infecção altera a temperatura por meio da agitação molecular, meus pensamentos movem-me os dedos através de uma longa cadeia de impulsos neurais do sistema nervoso até as fibras musculares. Esta regra vige para os vários reinos das ciências particulares, assim formando uma longa escalada hierárquica entre propriedades de níveis diferentes com o mundo físico fechado garantindo a base do conjunto.

O PFC sustenta a possibilidade de completude teórica (idealizada) de uma descrição fisicalista do mundo, nas quais verdades a respeito da dinâmica causal deste mundo formassem uma longa conjunção. Chalmers admite esse princípio sem maiores dificuldades. Ao contrário, a teoria dele depende da assunção da completude e mesmo do seu caráter apriorístico⁶.

Diante do dilema, têm-se duas hipóteses: ocorrências mentais causam efeitos mentais diretamente, (M causa M') ou causam efeitos físicos, que por sua vez, causam efeitos mentais (M causa P' que é a conjunto-base de M'). Ambas as opções mostram que há concorrência de causas, porque para cada uma

⁶ Isso fica claro em *Conceptual Analysis and Reductive Explanation*.

dessas causas mentais há uma causa física P anterior que se liga a P' (e P' é conjunto base de M').

A concorrência de causas, por seu turno, inflige o princípio da exclusão causal (PEC)

(PEC): Se uma propriedade X é suficiente para a causa de um evento Y, então não há nenhuma propriedade X' que seja causalmente relevante para o evento Y

É importante sublinhar que é pressuposto tanto no princípio do fechamento quanto no da exclusão causal, a ausência de mediação entre uma causa suficiente e seu efeito. Supondo que não houvesse, seria possível tecer cadeias com causas mentais intermediárias do seguinte modo: um efeito físico A, provoca um pensamento B, que por sua vez causa um efeito físico C. Entretanto, na formulação estrita de Papineau, todo evento tem uma causa física suficiente e imediata (PAPINEAU, 2008, p.57-60). Diante disso, proponentes de teorias não reducionistas tentaram solucionar o problema, entre outros meios, ao apelar para a noção de sobre-determinação.

Em termos gerais, a sobre-determinação se dá quando, em concorrência causal, duas ou mais causas são suficientes para produzir aquele efeito. Um exemplo típico de sobre-determinação é o seguinte: imagine-se um homicida que tivesse envenenado sua vítima e, para assegurar sua morte, tivesse lhe dado um tiro. Coincidentemente, no momento em que a bala atinge o crânio do homem partindo-o, o veneno paralisa o seu coração. Apresentam-se, neste quadro, duas causas concomitantes, as quais, separadas, provocariam o mesmo efeito, a saber, a morte da vítima. Neste caso, pode-se dizer que a bala não foi causalmente relevante para o acontecimento? Foi. Contudo, outra causa associa-se a ela para produzir o efeito, ainda que hipoteticamente este se desse mediante apenas uma delas (SPERRING, 2004, p.3-4).

A noção de sobre-determinação é uma saída. Mas há outras tentativas de solucionar a antinomia. A argumentação reducionista encontra dificuldades em relação a esquemas de eventos que possuem descrições complexas como

os que (DAVIDSON, 1970, pp. 214) apresenta em sua própria noção de sobreveniência:

"Embora a posição que eu descrevo negue que haja leis psicofísicas, ela é consistente com a visão de que características mentais são em algum sentido dependentes de, ou sobrevenientes a, características físicas. Tal sobreveniência pode ser interpretada como significando que não pode haver dois eventos similares em todos os aspectos físicos, mas diferentes em algum mental, ou que um objeto não pode se alterar em algum aspecto mental sem o fazer também em relação a algum físico."

DAVIDSON (1970, p. 208-209) expõe três princípios no início de "Mental Events": a) O da interação causal, estruturado em torno da ideia de interagir o mental causalmente com o físico (Davidson exemplifica com o modelo clássico da percepção. Temos a visão de um navio; esse é um evento sujeito à descrição mental, causado, contudo, por um input do ambiente externo - a presença do objeto extramental no campo visual). O modo como a percepção constitui-se na sua interação com o objeto é causal b) O já citado princípio do caráter nomológico da causalidade, que estabelece serem interações causais regidas por leis. c) E, finalmente, o princípio do anomalismo do mental, de acordo com o qual este não se subsume a leis psicofísicas estritas.

Davidson (1970, p. 209) formula a noção de anomalismo do mental tentando preservar uma intuição kantiana, à revelia do princípio do caráter nomológico da causalidade: a de que a filosofia não pode elidir a liberdade humana. Esta liberdade fundamenta-se na capacidade de autodeterminação da vontade, que vincula uma ação a um juízo decisório. Se a mente não causa nenhum efeito físico, ações no mundo, as quais inegavelmente são físicas, não são causadas por pensamentos. Assim, aparentemente, esvazia-se a autodeterminação da vontade, já que a vontade não poder de determinação sobre os atos (e, em última instância, nem mesmo sobre os pensamentos, já que eles podem ser reduzidos a disparos neurais).

A última citação de Chalmers na seção anterior revela uma posição convergente com o anomalismo do mental no tocante ao fechamento nomológico. Em *b*, o caráter nomológico da causalidade é assegurado. Para Chalmers, ele é assegurado e forma um sistema fechado.

A posição de Davidson foi matizada pela ênfase na distinção entre leis psicofísicas e leis psicofísicas estritas. Como resultado, ele buscava preservar esta autonomia supracitada, essencial a uma noção de agência moral, a qual, por sua vez, lastreia-se na formulação de juízos fenomênicos de autocompreensão.

Isto se torna possível sem ferir o princípio do fechamento, o caráter nomológico do mental ou o princípio da exclusão causal, em virtude da definição de evento que Davidson fornece. Para ele eventos caem sob descrições múltiplas. Consequentemente, um evento mental também é físico em outras descrições, muito embora constitutivamente físico em sua base ontológica. À diferença dos demais eventos físicos, ele possui uma semântica especial, que inclui verbos mentais.

Deste modo, a multiplicidade dos envoltórios semânticos, por assim dizer, importa em que o mesmo evento possui tanto propriedades físicas (sob descrição física) quanto propriedades mentais (sob descrição mental). Se a cadeia causal liga eventos uns aos outros, e não diretamente propriedades - uma abstração lógica - o resultado é que eventos com descrições adicionais de natureza mental são causalmente relevantes para outros *per se*, já que a causalidade física está imbricada nos mesmos.

Neste passo, resta saber se as descrições mentais dos eventos sofrem alguma espécie de isolamento causal, de modo a que as propriedades mentais não tenham eficácia, mas apenas as propriedades físicas. Tanto a solução a partir da sobredeterminação quanto aquela a partir das descrições múltiplas de Davidson são aportes clássicos para a posição não reducionista a respeito da causação do mental.

Não obstante, como visto anteriormente, o projeto de Chalmers parece conviver bem com o fechamento dos sistemas físicos. Ainda assim, as consequências contraintuitivas são suficientemente agudas para que se coloque um problema.

2.4 CIÊNCIA NÃO REDUCIONISTA DA CONSCIÊNCIA E ANTINOMIA CAUSAL

Chalmers aponta para a necessidade de construção de uma ciência não reducionista da consciência. Para realizar tal ciência, Chalmers estabelece a necessidade de articulação entre os dados obtidos a partir da perspectiva em primeira pessoa e aqueles obtidos a partir da perspectiva em terceira pessoa. Os primeiros são veiculados através dos 'relatórios verbais' (verbal report), fornecidos por um sujeito consciente, os quais informam um observador externo das experiências subjetivas que se passaram no âmbito interior do falante. Já os segundos são informações oriundas de exame, pelo observador externo, dos padrões cerebrais, comportamento, interações causais e demais aspectos independentes da subjetividade do examinado.

Há uma tendência na filosofia da mente contemporânea a articulá-los de forma unilateral: explicar os primeiros exclusivamente em termos de informações concernentes à dinâmica e função cerebral. Contudo, se estiver correta a premissa de que fatos relativos à estrutura e dinâmica dos sistemas físicos são insuficientes para explicar dados de primeira pessoa, e se quisermos preservar a legitimidade do *hard problem* da consciência, deve-se rejeitar tal hipótese explanatória.

Alguns dados centrais de primeira pessoa Chalmers menciona, de cuja articulação a ciência subjetiva da consciência depende,

- experiências visuais
- outras experiências perceptivas
- experiências corporais
- imaginação mental
- experiências emocionais
- pensamentos

Nesta lista, destaca-se o último, que, mais de uma vez, foi considerado a chave teórica do problema da consciência. Neste, se incluem decisões de natureza prática, cuja finalidade cumpre-se na realização de determinadas ações.

A dificuldade de compatibilizar os princípios do fechamento causal, da completude e da causalção mental, leva boa parte das teorias não reducionistas a um impasse. Esse impasse normalmente tem sido resolvido por meio da rejeição de algum dos princípios, frequentemente o do fechamento causal. No entanto, não é o caso com o dualismo elaborado por Chalmers. Nele, a completude do sistema total da física e o fechamento causal são ambos afirmados; a despeito disso, há lugar para uma sofisticada teoria do aspecto duplo da informação, que unifica o tratamento de sua realização física e fenomenológica. As consequências contraintuitivas porventura decorrentes dessa posição para a questão da causalidade mental não parecem gerar tensões teóricas decisivas dentro do sistema.

Para ilustrar este último ponto, mencionaremos uma conhecida e, sob múltiplos aspectos, desastrosa resenha de *The Conscious Mind* (SEARLE, 1997, pp. 135-176). Nesta review, com notável agressividade, Searle critica os resultados contraintuitivos a que supostamente chega a abordagem de Chalmers, tornando incorretas ou meramente analógicas proposições como "A dor no meu dedo foi provocada por uma martelada" ou "Vou pegar o lápis porque quero". Chalmers contra-argumenta de modo bastante esclarecedor. Peço desculpas pela longa citação, mas ante a acusação de eliminar a relevância causal da consciência sobre as ações físicas, ele diz:

Primeiramente, eu não digo que a consciência seja causalmente irrelevante para a ação. Esta é uma questão acerca de que sou neutro, e penso que há pontos de vista interessantes que dão à consciência uma relevância causal sem quaisquer danos à visão de mundo científica. (Por exemplo, fazendo que a consciência corresponda ao aspecto intrínseco de estados físicos que físicos caracterizam apenas extrinsecamente).

Em segundo lugar, eu não digo que a consciência nunca possa ser usada para explicar a ação, de modo que explicações que envolvam a consciência sejam inválidas. Eu simplesmente digo que invocar a consciência não é *necessário* para explicar ações; sempre haverá uma explicação física que não invoque ou implique a consciência. Uma melhor frase teria sido "explanatoriamente supérflua", ao invés de "explanatoriamente irrelevante". Algo pode ser supérfluo e, ainda assim, ser relevante.

Em terceiro lugar, não é verdade que minha visão implique que "se você acha que você está lendo porque você conscientemente quer ler, você está enganado". Em minha visão, é muito provável que você esteja lendo porque você queira ler. É só que o fato de que o desejo de ler é conscientemente experimentado não é necessário para a explicação se dar. Desejos conscientes podem explicar ações, e desejos inconscientes também podem explicar ações. Similarmente, você bebe porque está com sede, mas o aspecto conscientemente experimentado de tal sede não precisa ser considerado para explicar sua ação.

É claro que permanece algo contra-intuitivo quanto à superfluidade explanatória da consciência. Mas eu creio que se é forçado a isso por bons argumentos. E, no livro, eu

defendo isso extensivamente, argumentando que seja meramente contra-intuitivo e não apresente falhas fatais. (CHALMERS, Reply to Searle: Disponível em: <http://consc.net/books/searle-response2.html>)

Desde logo, Chalmers parece interpretar de forma singular a suficiência das causas físicas. Como se viu, anteriormente, existe certo consenso em considerar que eventos cujas causas físicas são suficientes prescindem de causas de outra natureza (excetuada eventual sobredeterminação), e que isto reforça o argumento contra a causação mental. Todavia, Chalmers posiciona-se de maneira neutra em face desse problema. Searle o interpreta, ao contrário, como se as suas teses implicassem uma tomada de partido no mesmo sentido da argumentação de Kim, usual entre materialistas reducionistas, mas incomum entre proponentes de teorias não reducionistas, materialistas ou dualistas.

De acordo com a interpretação de Searle, a atribuição causal de ações a decisões, constituindo parte da autocompreensão humana, pareceria dever ser tomada como expressão de um tipo de folk psychology dentro do projeto de Chalmers. Reforça-o o fato de que, ao invés de postular uma interação causal, Chalmers apela a princípios de coerenciação entre dados fenomênicos e seus vários tipos de correlato neuronal. Em busca de uma ciência da consciência, ele procura estabelecer princípios psicofísicos para explicar a correspondência verificada empiricamente entre estados fenomênicos e funcionais. Observa-se, por exemplo, a existência de uma associação entre a dinâmica da captação da luz, sua transformação em impulsos cerebrais e a experiência fenomênica da visão. Essa última, inacessível à metodologia das neurociências, pode ser constatada por meio dos relatórios verbais (*verbal reports*) em primeira pessoa, os quais guardam razoável grau de confiabilidade.

Visando a explicar esse tipo de associação, ele postula dois princípios: o princípio da coerência estrutural e o princípio da invariância organizacional. O princípio da coerência estrutural (principle of structural coherence) (CHALMERS, 1996, pp. 222-225) explica a correlação que habitualmente se verifica entre dados fenomênicos e interações cerebrais. Um dos traços dessa coerência é observado nos julgamentos fenomênicos. Estes julgamentos são notavelmente corretos: quando julgamos que temos uma sensação visual geralmente a temos; quando afirmamos que sentimos dor geralmente a sentimos. A este aspecto de

segurança nos juízos de segunda ordem, Chalmers denomina princípio da confiabilidade (reliability), o qual ampara a metodologia de investigação dos dados de primeira pessoa⁷.

O princípio da coerência estrutural estabelece a correspondência dinâmica entre os campos da consciência e o da awareness (traduzida, grosso modo, como percepção), esta última entendida como a disposição da informação para acessibilidade potencial e reportabilidade, uma noção dada em termos puramente funcionais. Pode-se observar esta relação na experiência visual.

Em uma experiência visual existe uma detalhada geometria de formas, profundidade, bem como uma estrutura implícita para informações fenomênicas de cor, tom, matiz, brilho, entre outras. Todas essas informações são disponíveis para controle do comportamento e reportabilidade dentro de um campo perceptivo – a noção de awareness traduz esse ‘campo’ – de modo que a experiência não é um dado difuso. Modificações nesse campo covariam com os estados fenomênicos qualitativos associados a cada uma dessas informações.

Já o princípio da invariância organizacional (principle of organizational invariance) estabelece que uma mesma estrutura material funcional possui os mesmos correlatos fenomênicos. Se dois padrões causais neuronais fossem reproduzidos em uma base biológica e um molde de silicone, haveria uma mesma experiência qualitativa associada a cada um deles. Chalmers subscreve a opinião bastante difundida de que a teoria da identidade clássica está equivocada porque estados mentais admitem múltiplas realizações. Se a organização funcional das múltiplas realizações for o fator decisivo e único para a emergência das propriedades fenomênicas, então o princípio da invariância organizacional é uma espécie de corolário das múltiplas realizações do mental.

Todavia, percebemos que a coerência genérica entre o mental e o físico salvaguardada por esses princípios, associados ao fechamento causal não são aptos a recuperar uma noção forte de agência, de ato intencional, de ato moral e de autodeterminação da vontade.

⁷ A autoridade dos juízos de segunda ordem não é absoluta. Muitas vezes, quando julgamos que temos dor confundimos categorias sensoriais. Ainda assim, há razoável grau de segurança nesse tipo de juízo.

Com efeito, Chalmers nega que as explicações por meio de propriedades fenomênicas sejam irrelevantes. Proposições do tipo listado por Searle não são no projeto deste autor, simplesmente, equivocadas. Resta certa dificuldade em entender como essa superfluidade das explicações fenomênicas não incide na mesma dificuldade que a irrelevância das causas mentais. Se as explicações que apelam a propriedades físicas são suficientes para explicar todos os eventos, então explicações fenomênicas pareceriam ser mantidas apenas como um hábito supérfluo. Não existe nenhuma tentativa séria de articular a existência dessas explicações como o fundamento dos juízos fenomênicos que surge como sendo a chave para a questão da consciência.

|

CAPÍTULO 3: POR UMA CIÊNCIA DA CONSCIÊNCIA

Cumpramos expor alguns dados científicos relevantes, trazidos por David Chalmers e Peter Carruthers com o intuito de fundamentar a construção de uma ciência da consciência. Tentaremos também oferecer uma conciliação entre as perspectivas defendidas por ambos os autores no nível da descrição científica das operações da consciência de acesso (*access consciousness*).

Chalmers e Carruthers têm diversos pontos de convergência e divergência. Acredito que o modo como eles interpretam os dados científicos é um forte ponto de convergência. Em eventuais cruzamentos teóricos, caberia apenas enfrentar questões de natureza conceitual, o que simplificaria notavelmente a tarefa de aproximar dois dos mais importantes proponentes de uma ciência unificada da consciência.

Desde logo, uma das distinções mais usuais realizada por filósofos da mente diferencia dois tipos de consciência: a fenomênica e a de acesso (*Access consciousness*) (BLOCK, 1995, pp.227-2288). O primeiro tipo corresponde às experiências que um agente sente de modo subjetivo - cores, sons, sensações provocadas internas e externamente, o componente qualitativo dos conceitos fenomênicos. A famosa pergunta de Thomas Nagel - *what is it like to be a bat?* - (NAGEL, 1974, pp. 435-440) indica que, nesta instância, existe um "*like to be*", apreendido por um ato de transparência cognitiva de primeira pessoa. Dizer de alguém 'O sujeito S é consciente' consiste em atribuir-lhe uma experiência qualitativa ou julgá-lo potencialmente capaz de possuí-la; dizer que tal estado mental é consciente, por sua vez, será verdadeiro caso o estado mental compartilhe esse traço qualitativo.

Já a consciência de acesso corresponde a um conceito funcional. Estados conscientes, neste sentido, seriam definidos como estados global ou extensivamente disponíveis para sistemas cognitivos, a exemplo de sistemas de formação de memórias, inferenciais, afetivos. Nesta acepção, é consciente o estado que se mostra acessível a essas diversas funções. Se o estado mental expresso no juízo "Está chovendo" pode compor inferências tais como "Está chovendo, então não vou sair de casa"; ou disparar gatilhos emocionais como

em "Está chovendo. A chuva me deixa melancólico"; ou, ainda, ser o conteúdo de uma memória "Está chovendo" (lembrança de uma cena passada), isto se deve ao fato de que *tal estado é consciente*. Um estado inconsciente não seria acessível a todos esses sistemas, embora possa influenciar, em alguns casos, o comportamento do agente.

Afirmam tanto (CHALMERS, 1996; 2010) quanto Carruthers que os dois tipos de consciência se relacionam de um modo estreito. Nota-se, por exemplo, que para (CARRUTHERS, 2000) um estado consciente, cuja existência pode ser acessada pelo sujeito, é candidato a ser explicado, no plano conceitual, por teorias representacionistas (HOR, HOD). Já no plano empírico, são fundamentais as explicações científicas que isolam a consciência de acesso e mapeiam a interconexão entre as áreas cerebrais ativadas. Os estados fenomênicos também podem ser explicados em função dos mesmos experimentos. Assim, a sustentação empírica de uma ciência da consciência abrigará, especialmente, trabalhos que correlacionem os fatores físicos responsáveis pela consciência de acesso e capazes de explicar a arquitetura de funções como atenção e memória de trabalho, essenciais para a acessibilidade dos estados mentais. (CARRUTHERS, 2015 pp.75-107).

Também (CHALMERS, 1996, p.238) argumenta em sentido convergente. Ele postula um princípio de coerência estrutural entre consciência fenomênica e *access consciousness (awareness)*, que deve imperar como prescrição metodológica para a interpretação dos dados empíricos obtidos através do exame dos mecanismos neurais de ativação da consciência de acesso. O princípio da coerência estrutural, conforme apontado anteriormente, indica uma correlação entre o aspecto fenomênico da consciência e a base natural, a cuja materialidade ela sobrevém. O que (CHALMERS, 2000, p.59) entende por este correlato natural não é uma noção tão simples.

Em primeiro lugar, ele indica uma definição provisória de correlato neural. Nesta definição, ele seria

Um sistema neural N é um NCC (Correlato Neural da Consciência) se o estado de N correlaciona-se diretamente com estados da consciência. (CHALMERS, 2010, p. 61)

Ele avalia que o correlato neural da consciência deve ser encontrado nos fatores que tornam certos estados mentais disponíveis a ela. Mas a sua definição será refinada em razão de alguns aspectos. Primariamente, a consciência possui dois níveis, a “consciência de fundo” – cuja descrição é a de um estado como estar acordado, dormindo, sob hipnose, dentre outros – e os conteúdos da consciência. Quando a determinação do NCC diz respeito à consciência de fundo, trata-se de verificar a correlação entre um estado geral do indivíduo, dentro do qual conteúdos conscientes específicos se dão, e o correlato neural. Para esse último, parece ser o caso de se encontrar um correspondente nos sistemas talamocorticais cruzados com níveis neuroquímicos de ativação.

Diferente é a avaliação da correlação neural para os estados conscientes, isto é, aqueles estados que aparecem contrastados à consciência de fundo como possuidores de um conteúdo específico. Tais estados incluem a experiência de uma imagem em particular, de um padrão sonoro ou de um pensamento consciente. Estes estados são especificados pelo seu conteúdo. Assim, especifica-se o estado cuja descrição é “ter a experiência de uma superfície vermelha”. Ele é especificado por este preciso conteúdo: aquele estado é o de “ter a experiência de uma superfície vermelha”. Ora, não é seguro supor que tenham as mesmas características os NCCs da consciência de fundo e os NCCs dos estados individuados da consciência; afinal, os correlatos neurais para quem está tendo a visão de uma superfície vermelha não são os de quem simplesmente está acordado vendo *algo*. O que se pretende determinar é qual o correlato para a experiência de uma curva, do vermelho, de um dó7 na escala e assim por diante. Desta forma, segundo a tese citada por (CHALMERS, 2010, p.64), os correlatos neurais destes estados individuados concernem aos aspectos representacionais das experiências.

Chalmers refina o conceito imensamente, adicionando várias especificações a partir dessa segunda definição. Uma especificação de notável importância, relacionada estreitamente à pesquisa empírica, é que o NCC deve ser *suficiente* (o uso feito por Carruthers da pesquisa empírica em matéria de ciências cognitivas dá sinal disso). Essa especificação é importante, posto que

qualquer mapeamento em fMRI mostra que o cérebro está em atividade em múltiplas áreas distintas, com interconexões complexas por todo o sistema nervoso em um dado tempo. Um dos desafios do cientista é justamente elaborar experimentos que lhe permitam isolar aquela atividade cerebral específica cuja presença explica um fenômeno. Do contrário, ao não se poder isolá-la, não haverá critério para assinalar aquela atividade ao fenômeno consciente. Ao estabelecer a condição de suficiência, a abordagem de Chalmers coaduna-se com o tipo de estudo mais frequentemente citado por Carruthers. Ao final do artigo, ele chega à sua definição acabada

Um NCC é um sistema neural mínimo N, tal que existe um mapeamento de estados de N para estados de consciência, em que a um dado estado de N é suficiente, sob condições C, para o correspondente estado de consciência

(CHALMERS, 2010, p.91-100) dedica-se a outra especulação: se ele antes queria dizer o *que* era o NCC, ele pretende então dizer *como* encontrá-lo. Esta seção é também de particular relevância para se perceber o quão profundas são as afinidades entre os dois autores. A essa prévia abordagem, anterior ao trabalho de experimentação científica, ele denomina de busca por *pre-experimental bridging principles*. E, basicamente, refere-se a duas questões: a) se um sistema é ou não consciente b) quais informações contidas neste sistema são conscientes e quais não o são.

Coerente com a tese não-reducionista acerca das propriedades conscientes, ele adverte que esses princípios não são diretamente testáveis. Não se pode saber se um sistema é consciente ou não a partir de dados científicos relativos à sua estrutura neural, organização funcional ou qualquer aspecto analisado de maneira empírica. Faz-se possível apenas obter inferências dos resultados científicos para as conclusões válidas sobre *a* e *b*. Tais inferências, contudo, não podem possuir um caráter de necessidade lógica. Afinal, se o observador científico pudesse inferir, dos dados empíricos de um sistema, que ele *necessariamente* é consciente (em sentido fenomênico), ou que tais informações são *necessariamente* conscientes, então a distinção entre sobreveniência lógica e natural seria prejudicada. Essa necessidade seria, pois,

um indicativo de que a sobreveniência da consciência no cérebro é lógica e, por isso, necessariamente aquele sistema é consciente.

O que fica claro é que estes princípios têm caráter interpretativo. Da parte do observador, tomam-se certos dados para que, ao interpretá-los como indicativos da consciência do sistema ou de uma informação nele, possa se estabelecer a inferência. A inferência estabelecida não é uma relação de necessidade lógica. Ora, na interpretação destes dados destaca-se, especialmente, a reportabilidade verbal, abordada em capítulos anteriores. Deste modo, quando se quer saber se determinado agente é consciente ou se tal informação o é, um dos modos de se ter acesso a isso passa por questionar o agente sobre as suas experiências atuais e potenciais.

Um dos problemas com este critério é o fato de ele automaticamente excluir agentes que não possam dar relatos verbais: todo o vasto mundo da consciência animal, por exemplo, está fora dele. Para contornar essa dificuldade, experimentações com animais costumam isolar um aspecto de consciência fenomênica em uma relação estímulo-resposta. Desta forma, se um macaco é posto em um experimento controlado no qual ele precisa desempenhar certa tarefa quando um estímulo consciente o move, saberemos que essa informação é consciente naquele sistema neural – o do símio. Se ele deve apertar um botão quando vê uma cor vermelha aparecer em um painel; e, se, de outro modo, não o faria, caso o faça, então restará claro que a informação da cor do painel é consciente para o macaco. Cumpre notar que o modo de contornar a dificuldade ilumina um aspecto central para Chalmers e Carruthers: a disposição para o controle global que caracteriza a consciência. A informação consciente está passível de ser utilizada por outras funções, a exemplo da motora, e isso caracterizaria a consciência pacificamente em ambos os autores, ao menos em seu aspecto psicológico (CARRUTHERS, 2015, p.48-54).

Chalmers aprofunda esta interpretação ao estabelecer uma reconstrução racional nestes termos,

1. Consciência ↔ disponibilidade global (bridging principle)
2. Disponibilidade global ↔ processo neural (pesquisa empírica)

3. Consciência ↔ processo neural N (conclusão)

Essa estrutura pode ser exemplificada: se certos estados mentais dependem de uma oscilação a 40-hertz para serem globalmente disponíveis, então, mediante o princípio da coerência estrutural, o correlato neural referente ao aspecto fenomênico daqueles estados se encontra ali. Sob este aspecto, a avaliação de Chalmers é metodologicamente compatível com a perspectiva desenvolvida por Carruthers. Há duas importantes diferenças entre elas, contudo.

1) A irreducibilidade ontológica e inter-teórica do aspecto fenomênico da consciência é afirmada por (CHALMERS, 1996; 2010), ao passo que (CARRUTHERS, 2000, p.59) a nega. Em consequência disso,

2) A investigação científica sobre os estados conscientes orienta-se para a determinação de um Correlato Neural da Consciência (CHALMERS, 2010, p.59), enquanto a abordagem adotada por Carruthers elide a questão de um correlato neural. Ao contrário, argumenta-se a favor de uma versão da tese da identidade, assim como de uma explicação representacionista dos estados fenomênicos - HOR (High Order Representationalism) (CARRUTHERS, 2000; 2015).

Tanto a exposição de Chalmers quanto a de Carruthers sobre a noção de consciência de acesso são tributárias da teoria do espaço global (BAARS, 1988). A teoria tem acumulado grande número de evidências a seu favor, de modo a gozar de amplo acolhimento na comunidade científica. Neste passo vai a reconstrução de Carruthers, inteiramente tributária de um amplo número de estudos científicos.

Baars postula uma mudança de estado mental inconsciente/consciente que se dá mediante acréscimo no nível de ativação neuronal. Este *upgrade* coaduna-se com a arquitetura neural do cérebro e também recebe forte suporte experimental (SERGENT & DEHAENE, 2004, pp. 720-727). Dois experimentos foram conduzidos e indicaram que o paradigma gradualista da consciência não dá conta das evidências. Eles mostram que, a partir de certo ponto de

interconexão entre áreas afastadas do sistema neural, um determinado objeto é percebido visualmente, corroborando os pressupostos da arquitetura do espaço global.

Uma versão mais recente da teoria do espaço global é a teoria do Global Neuronal Workspace, introduzida por (DEHAENE & CHANGEUX, 2011, p.210). Este sistema de processamento da consciência-acessibilidade pode ser visto como uma variante da teoria do Espaço Global de Baars. Basicamente, a GWN afirma a presença de uma estrutura neural de longo alcance, densa nas áreas pré-frontal, cingulada e parietal que, junto com os loops talamocorticais relevantes seria responsável por conectar e sustentar a atividade de processos conscientes no GWN.

O artigo faz uma extensa revisão da literatura científica para prová-lo. Os resultados convergem com a interpretação hipotética de (CHALMERS, 2010, p.91). Em “On The Search for the Neural Correlate of Consciousness”, Chalmers propõe que, se houver uma área do cérebro que faz uma ponte entre a área pré-frontal e o córtex visual, então deve residir aí o correlato neural da consciência visual. Como se vê, essa hipótese é confirmada pelo artigo (DEHAENE & CHANGEUX, 2011). Mais ainda: ele afirma serem as idéias de Baars compatíveis com a sua perspectiva teórica, embora não implicadas por ela.

Estas evidências suportam-se mutuamente e confirmam a intuição de que processos neurais nas áreas pré-frontal e parietal estão intimamente associadas à acessibilidade em um espaço global, e que a atenção guarda especial papel neste modelo.

Carruthers pretende utilizar este framework para fornecer uma teoria de como o correlato neural explica o funcionamento da consciência de acesso. O papel da atenção e da memória de trabalho destaca-se especialmente nessa teoria.

Carruthers (2015, pp.54-56) afirma haver ampla evidência de que o espaço global, e, portanto, a consciência de acesso, seja dependente da atenção. Por exemplo, no artigo de (MACK e ROCK, 1998, pp.11-17) um experimento com cerca de 5000 pessoas é descrito. Resumidamente, uma cruz aparece na parafovéa (região relativamente central do arco da visão) e uma

imagem mascarada surge. Pelas cobaias que mantiveram a atenção fixa na cruz essa última imagem não foi notada. Esse resultado levou os pesquisadores a mudarem radicalmente a assunção inicial da pesquisa e a investigarem, em consequência disso, o fenômeno da Inattentional Blindness (IB). Evidências desta natureza sugerem que a atenção é um elemento necessário para a percepção consciente, embora não seja suficiente (CARRUTHERS, 2015, p.57-58). Mas, se a atenção é necessária para a consciência de acesso, como ela funciona?

Carruthers enuncia esta explicação para o funcionamento da atenção: ela aumenta a sensibilidade de um conjunto de neurônios e suprime a atividade de outros conjuntos. Por exemplo, a atenção voltada a formas espaciais aumenta a atividade de grupos específicos de neurônios na área V4, ao passo que a atenção voltada a movimentos apresenta um incremento na área MT.

Reynolds e Chelazzi (2004, pp. 611-47) apoiam-se em uma literatura científica já existente. Experimentos reiterados com humanos mostram que a atenção orientada a um estímulo pode torná-lo perceptível; além disso, através desse foco de atenção é possível captar propriedades do mesmo estímulo que, de outro modo, não seriam notadas. Como argumentam os autores, estudos recentes em psicologia experimental com mapeamento das áreas cerebrais mostram que um acréscimo de atividade nas áreas responsáveis pela catalisação da atenção implicam saltos de percepção consciente em um paradigma descontínuo.

Essas são algumas fontes científicas citadas no extensíssimo trabalho de interpretação científica conduzido por Carruthers. Cumpre tornar clara a existência de uma orientação nítida dele no sentido de se apoiar em estudos que enfatizam o papel da disponibilidade global e da atenção para o fenômeno da consciência. Se é assim, então podemos construir o seguinte esquema de raciocínio, verdadeiro tanto para Carruthers quanto para Chalmers, exceto pela *interpretação* da função da disponibilidade global.

1. Consciência ↔ disponibilidade global (bridgingprinciple)
2. Disponibilidade global ↔ processo neural (pesquisa empírica)

3. Consciência ↔ processo neural N (conclusão)

A diferença não se dá nos elementos das premissas e na conclusão, mas, sim, na interpretação da primeira premissa. Ao invés de considerar que o bicondicional “Consciência ↔ disponibilidade global” é uma ponte interpretativa ligando o processo neural e a consciência, Carruthers aborda-o ontologicamente; a disponibilidade global é uma característica da natureza dos estados conscientes porque as propriedades conscientes são as próprias propriedades físicas.

3.1 DISCORDÂNCIAS FUNDAMENTAIS

3.1.1 Posicionamento de Carruthers: estratégia dos conceitos fenomênicos

As similaridades entre as abordagens de Chalmers e Carruthers quanto à interpretação dos dados científicos é um ponto suficientemente bem estabelecido. Todavia, as implicações filosóficas extraídas por ambos desses dados são diversas em razão de concepções gerais distintas trazidas por cada um deles sobre a ontologia e a epistemologia da mente. Basicamente, os aspectos mais sensíveis do problema situam-se na aceitação ou rejeição dos dois reducionismos típicos - ontológico e epistemológico. Enquanto Chalmers não reduz as propriedades fenomênicas a propriedades físicas, seja ontológica ou epistemologicamente, Carruthers o faz em ambas as dimensões.

Não obstante, *Phenomenal Consciousness* fornece uma imagem a um só tempo elaborada e extensiva da argumentação acerca do tema que os divide. Por este motivo, os argumentos deste livro serão considerados como exibindo as posições corretas do autor, excetuadas possíveis reavaliações posteriores. Carruthers se põe em acordo com Chalmers quanto à descrição do processo de redução *explanatória*. Esse tipo de redução acontece em fenômenos como a reprodução, a vida, a aprendizagem, dentre outros. Ocorre quando se pode

oferecer uma explicação dos fenômenos de alto nível por meio de processos de baixo nível. Deste modo, alguém que desejasse explicar a reprodução (fenômeno de alto nível) descrevê-la-ia por meio de termos e descrições extraídos da genética. Assim, a reprodução seria explicada como uma troca genética que se dá de um modo tal (descrição do processo de baixo nível da genética) e com isso estaria suficientemente explicada. A reprodução tem naqueles processos a sua explicação suficiente.

Contudo, essa forma de redução precisa ser devidamente diferenciada da redução inter-teórica. Ao menos em um sentido da última expressão, as reduções explanatórias não a exigem. Este sentido é o da clássica redução inter-teórica metafísica, que supõe a redução ontológica dos fenômenos de alto nível aos de baixo. Reduções inter-teóricas de sucesso se dão quando há uma *identificação* envolvida. Muito raramente isto acontece, por conta da múltipla realização das propriedades de alto nível, frequente nos campos da biologia e da psicologia, por exemplo. Deste modo, tomemos um fenômeno como a aprendizagem: ela pode ser explanatoriamente reduzida a operações funcionais de *input/output*, mas não pode ser *identificada* com aquele processo específico em que se realiza em um indivíduo concreto. Afinal, são vários os tipos de seres capazes de aprender; várias as suas conformações cerebrais específicas, bem como as suas leis de recepção dos estímulos e os processos informacionais. Portanto, caso a redução explanatória exigisse a inter-teórica, em todas as situações onde a identificação não fosse estrita, por conta da múltipla realização, seria impossível haver a primeira das reduções. Uma vez que decididamente ela é possível, então não exige a segunda.

Carruthers compartilha das mesmas premissas, rejeitando, contudo, as implicações filosóficas que Chalmers extrai delas. Ele prega serem os conceitos fenomênicos conceitos de reconhecimento. Um conceito de reconhecimento é um conceito aplicado nas instâncias em que há conhecimento por familiaridade, dependendo, de maneira evidente, da aceitação da distinção entre conhecimento proposicional e conhecimento por familiaridade. Diz-se conhecimento por familiaridade aquele em que se conhece *p* sem que este conhecimento dependa de qualquer inferência. Um conceito *puramente* de

reconhecimento aplica-se, pois, às instâncias do conhecimento por familiaridade, de modo isolado de qualquer outro conceito ou crença.

Isto significa dizer que, ao ter a experiência de vermelho, se o conceito envolvido no juízo "isto é vermelho" indicar puramente um modo de se reconhecer aquela experiência, então teremos a aplicação de um conceito de reconhecimento a um estado de experiência. Carruthers (2005, p. 83) defende um modo simples de se entender essa possibilidade. Ele concebe um indivíduo que, ao estar em um estado k , produz um julgamento de k que envolve um conceito de reconhecimento. O referido julgamento é provocado por um mecanismo causal. Mais ainda: é possível supor este mecanismo causal isolado funcional e causalmente de todos os demais mecanismos ao se ter um conceito puramente de reconhecimento da experiência. Conceitos fenomênicos frequentemente estão isolados causal, funcional e representacionalmente dos demais. Assim aparece no famoso experimento da cientista Mary no quarto preto-e-branco. Mary não pode deduzir, a partir de seu conhecimento total do cérebro, o conceito que envolve a experiência do 'vermelho' por que este é isolado dos outros.

Para reduzir o argumento de Chalmers ao absurdo, Carruthers propõe outro experimento mental. Imaginemos um ser que possuísse conceitos de quase-cor puramente de reconhecimento e não possuíssem qualquer experiência. Este tipo de zumbi de Carruthers é uma variante possível do experimento mental de Chalmers, uma vez que nele os zumbis realizam julgamentos fenomênicos embora não possuam experiência fenomênica.

Desta forma, os zumbis de Carruthers realizam julgamentos sobre cores. A grande e fundamental diferença é: os seus julgamentos são, por uma estranha razão, descolados daqueles feitos por observadores normais. Se os zumbis de Chalmers podem se passar por pessoas normais, porquanto seus julgamentos fenomênicos imitam os julgamentos de outros (a despeito da ausência de experiência qualitativa correspondente), os zumbis de Carruthers, por sua vez, produzem julgamentos com conceitos que não são compartilhados por observadores normais. Eles têm seus próprios conceitos de quasi-cor. Diante de uma superfície cujo estímulo lhes provoque a identificação como superfície quasi-vermelha, eles farão julgamentos aplicando seu conceito de quasi-

vermelho. Todavia, este não tem qualquer relação funcional, representacional ou causal com outros conceitos (ao contrário dos conceitos ordinários da linguagem das cores que o zumbi de Chalmers, mais bem inserido na sociedade, por assim dizer, utiliza).

Ora, disso ele extrai que, se os conceitos não têm relação causal, funcional ou representacional, então eles estão isolados. Reduzindo ao absurdo, conceitos de experiência como a representar um aspecto da experiência fenomênica irreduzível, são também puramente de reconhecimento. Portanto, o que experimentos mentais anti-reducionistas mostram é a existência de um isolamento de tais conceitos, implicando este a possibilidade de se conceber mundos possíveis onde julgamentos mediante eles são viáveis sem a experiência, não apresentando este fato, contudo, implicações metafísicas.

3.1.2 Resposta de Chalmers

Chalmers é consciente deste desafio imposto pelos materialistas. De fato, (CHALMERS, 2010, p.305) nota a existência de algumas respostas materialistas-padrão diante do anti-reducionismo epistemológico. Uma delas é acreditar que ainda não temos meios de reduzi-lo, mas que teremos um dia. Essa é uma posição de aposta na confiança da ciência, devidamente fundamentada por argumentos de probabilidade, disposta a provar a contingência do "epistemological gap". Outras respostas identificam o aspecto ontológico e o epistemológico; ao atacar o primeiro julgam vencer o segundo.

A estratégia que Chalmers enfrentará, porém, é outra. É a que Carruthers apresenta em *Phenomenal Consciousness*, denominada por ele de estratégia dos conceitos fenomênicos. Basicamente, ela consiste em dar uma boa explicação física sobre por que encontramos um gap explicativo. A estratégia dos conceitos fenomênicos mostra que o gap decorre exclusivamente do modo como pensamos a consciência e de não como ela é. Outra das consequências de sua adoção é a rejeição de intuições comuns aos partidários de argumentos do conhecimento (como o argumento de Mary no quarto preto-e-branco).

Chalmers aprecia a estratégia como sendo a mais eficiente e interessante resposta materialista. Todavia, ele a considera insustentável. A forma do argumento bicondicional, já citado no capítulo 2, afirma que P são as verdades microfísicas do mundo. Q, uma experiência qualitativa. A relação entre concepção e possibilidade metafísica também fica clara no argumento,

1. $P \& \sim Q$ é concebível.
2. Se $P \& \sim Q$ é concebível, $P \& \sim Q$ é metafisicamente possível.
3. Se $P \& \sim Q$ é metafisicamente possível, o materialismo é falso.

4. Materialismo é falso.

A estratégia dos conceitos fenomênicos funciona do seguinte modo. Proponentes colocam a tese C, à qual atribuem certos aspectos psicológicos. (1) Eles argumentam ser a tese C verdadeira. Depois, que (2) tal tese serviria pra explicar a situação epistêmica em relação à consciência e, finalmente, que (3) C pode ser explicada fisicamente. A conclusão é que a consciência pode ser explicada fisicamente.

Concretamente, substituindo a variável C, Carruthers aposta na primeira das três táticas filiadas à estratégia fenomênica, a saber, a tática dos conceitos de reconhecimento. As outras duas são a tática do papel conceitual distinto e a dos conceitos indexicais. Para um proponente da primeira, é preciso mostrar que os conceitos de reconhecimento constituem um tipo de aspecto-chave capaz de explicar a razão pela qual temos a ilusão de que a consciência não é corretamente capturada em termos reducionistas. Ou seja, substituindo a tese C pela de Carruthers resta que conceitos fenomênicos são conceitos de reconhecimento, explicando isto a nossa situação epistêmica, bem como a aparência do salto epistemológico, passível de ser explicado fisicamente.

Para refutar a tese geral, Chalmers propõe um argumento

1. C não é fisicamente explicável
ou
2. C não explica nossa situação epistêmica a respeito da consciência

A escolha entre essas duas opções é balizada pela seguinte pergunta: P&~C é concebível? Ou seja, podemos conceber seres idênticos a nós que não possuam a característica C (CHALMERS, 210, p. 312). Os zumbis funcionais, semelhantes aos zumbis clássicos de Chalmers, além de não possuírem consciência também não possuem os aspectos psicológicos envolvidos na tese C. Se pensássemos o exemplo de Carruthers, estes zumbis não teria nem consciência nem conceitos puramente de reconhecimento. O argumento que se constrói a partir disso é o seguinte:

1. Se P&~C é concebível, então C não é fisicamente explicável
 2. Se P&~C não é concebível, então C não pode explicar nossa condição epistêmica
-
3. Então C não é fisicamente explicável, ou C não pode explicar nossa situação epistêmica

A defesa da verdade da premissa 1 se dá por uma estratégia semelhante à estratégia original do argumento bidimensional em relação a P&~Q. A conexão que existe é de natureza epistêmica, não metafísica. P&~C, se concebível, C não pode ser explicada fisicamente por que a explicação que se busca é uma que torna transparente o elo entre verdades de baixo nível e verdades de alto.

Um proponente da estratégia através dos conceitos quasi-fenômênicos pode tentar argüir que não é necessário que C seja uma tese sobre conceitos fenomênicos. Ela pode ser uma tese sobre conceitos quase-fenomênicos. Essa tática é adota por Carruthers. Neste caso, o que seria fisicamente explicável seria a tesa C relativa aos conceitos quase-fenomênicos. Tais conceitos seriam

capazes de explicar a nossa condição epistêmica sem apelar para propriedades fenomênicas. E, uma vez que os conceitos quasi-fenomênicos são puramente de reconhecimento, a sua obtenção se dá através de conhecimento por familiaridade, diretamente, não mediado por inferências de e para outros conceitos.

Ora, a consequência disso é que se os conceitos quase-fenomênicos em C dependem de uma estrutura de causalidade direta, não mediada por outros fatores causais. Se é assim, então P&~C não pode ser concebido. Caso possa, então, há um reforço ainda maior à idéia de que C não explica a relação de transparência entre as verdades microfísicas e as verdades psicológicas. Do contrário, caímos na outra hipótese do dilema.

Esta segunda hipótese consiste em que P&~C não seja concebível. Neste caso, então, necessariamente, um zumbi (alguém que não possui consciência) satisfará C. Isso ocorre por que, haja vista ter a mesma estrutura física de sua cópia natural, o zumbi necessariamente apresentará aqueles aspectos psicológicos assinalados pela posse do conceito que C envolve. Dois sujeitos fisicamente idênticos terão os mesmos conceitos fenomênicos ou 'quasi-fenomênicos'.

Porém, ocorre um problema. Chalmers adverte que zumbis não compartilham a nossa situação epistêmica. Então, se isso é verdadeiro C não pode explicar a nossa situação epistêmica, já que se os zumbis satisfazem C e não compartilham nossa situação epistêmica, é por que C não satisfaz uma explicação de nossa situação epistêmica.

O núcleo dessa argumentação é a premissa que diz que os zumbis não compartilham de nossa situação epistêmica. Existem muitas razões para defendê-la. Em primeiro lugar, (CHALMERS, 2010, p.316-317) alega razoavelmente que a situação epistêmica de um indivíduo deve incluir o valor de verdade e o status epistêmico de suas crenças (se justificadas ou não, se cognitivamente relevantes ou não). Caso o zumbi tenha essas mesmas crenças, igualmente verdadeiras e justificadas, então ele compartilha a situação epistêmica de sua cópia. Assumindo que zumbis possuam crença (o que é necessário para a estratégia dos conceitos fenomênicos funcionar), segue que os valores de verdade e os graus e modos de justificação devem ser idênticos.

Alguns experimentos mentais reforçam a noção intuitiva de que esse não é o caso. Poderíamos imaginar um diálogo entre dois zumbis-filósofos (CHALMERS, 2010, p.317-318). O zumbi eliminativista argumentaria defendendo a tese de que não existe algo como uma consciência fenomênica. Já o realista diria que sim, existe. O eliminativista asseguraria que a consciência só existe enquanto conceito funcional e mais nenhum sentido, ao passo que o realista diria que existe algum sentido além do conceito funcional

Ora, a posição não-eliminativista (mesmo reducionista) aceita a explicação do realista como verdadeira. No mundo dos zumbis, porém, essa explicação seria claramente falsa, porquanto os zumbis não têm, de fato, consciência em nenhum outro sentido exceto o funcional. Se assim o é, então os valores de verdade das crenças dos zumbis são diferentes dos valores de verdade das crenças de suas cópias humanas que travassem o mesmo diálogo. Ou então os não-eliminativistas seriam obrigados a considerar que no mundo humano os valores de verdade dessas crenças são os mesmos que no mundo zumbi; assim, teriam de abandonar sua posição. Logo, zumbis e cópias humanas não compartilham a mesma condição epistêmica ou os não-eliminativistas são forçados a abandonar sua posição.

Imaginemos outro exemplo; no argumento do quarto preto-e-branco, Mary possui todos os conhecimentos sobre a estrutura do sistema nervoso, mas nunca teve a experiência de ver uma cor fora da paleta preto-branco. Sua gêmea, Mary zumbi, encontra-se na mesma situação da cópia: em um quarto-e-branco, ela acumulou conhecimentos sobre o cérebro. Suponha que ela saia do quarto e veja uma maçã vermelha. Que tipo de conhecimento ela obteve (CHALMERS, 2010, p.318)? Note-se que a condição epistêmica de Mary e Mary zumbi deve ser a mesma para que C tenha função explanatória, uma vez que tanto o zumbi quanto Mary satisfazem C. Se zumbis satisfazem C, e a condição epistêmica deles é distinta da nossa, então C não é suficiente para explicar a nossa condição epistêmica.

Certamente, ele adverte, algum conhecimento foi obtido pela Mary zumbi: ela pode ter obtido conhecimento indexical de que estava naquele estado em dado momento. Também pode ter obtido conhecimento da habilidade de perceber e classificar certo estado. Todavia, o conhecimento relevante

fenomênico ela não obteve. Logo, a conclusão é que sendo P&~C concebível ou não, C não pode explicar a nossa condição epistêmica. Portanto C não pode explicá-la em nenhum caso possível.

3.2 ALGUNS DADOS SOBRE A MEMÓRIA DE TRABALHO

Não obstante as diferenças conceituais que separam Carruthers de Chalmers, no que se refere ao papel da memória de trabalho e da atenção, as semelhanças entre as duas perspectivas são notáveis. Ambos consideram, justamente, a centralidade da atenção para a consciência, assim como indicam fontes científicas convergentes (teorias do espaço global e variantes) como sendo as melhores explicações naturalistas para o fenômeno da consciência. Isto significa que, até certo ponto, estamos justificados em considerar uma convergência preliminar entre as explicações que Carruthers fornece sobre a memória de trabalho e os princípios gerais que devem orientar a formação de uma ciência da consciência. Obviamente, na ausência de uma explicação naturalista da consciência dentro da obra de Chalmers, não cumpre prejudicar a questão em favor da opinião de Carruthers em tópicos a respeito dos quais o primeiro não se pronunciou. A questão que se oferece diz respeito tão somente a compatibilidade *em princípio* entre as duas teorias, compatibilidade que se apresenta como coerência hipotética dados os princípios de Chalmers e as inferências explanatório-científicas de Carruthers. O ponto em que essas inferências se chocam com aqueles princípios diz respeito aos problemas tratados no capítulo anterior.

A memória de trabalho pode ser definida como sendo um tipo de memória destinada a armazenar e combinar conteúdos com pouco tempo de retenção. Esta faculdade revela-se notavelmente importante no gerenciamento da atividade cognitiva, pois permite a retenção da informação para a realização de um vasto espectro de atividades - compreensão da linguagem, operações matemáticas, raciocínio. Revela-o a leitura de períodos cuja ambiguidade semântica dos termos iniciais exige a memória de trabalho, uma vez que o seu significado só é completamente saturado quando a última peça de informação é

encaixada. A mesma dependência da memória de trabalho acomete operações de cálculo, processos decisórios com cadeias de inferências.

A memória de trabalho é classificada junto com a de longo termo e sensorial, as quais são compreendidas, respectivamente, como a capacidade de retenção de informações por longos períodos e a capacidade de retenção de informação sensível em curtíssimo período. Forte evidência empírica permite distinguir entre a memória de longo prazo e a de curto prazo. Dentre os indícios disponíveis, destacam-se experimentos em pacientes que sofriam síndrome de Korsakoff e que, não obstante possuíssem amnésia retrógrada, eram capazes de memorizar sequências numéricas ou léxicas. Desde os anos 60 havia sido firmada a distinção entre memória de longo e curto prazo.

Os resultados de investigações científicas sobre a memória de trabalho mostram que ela é composta de três sistemas, um principal, executivo, e outros dois subordinados. O sistema executivo funciona como uma central de agenciamento e controle. Na teoria de Baddeley, ele utiliza sistemas de armazenamento colaterais também empregados na percepção e na fala. Ligados ao sistema executivo situam-se os dois subordinados. Um destes dois últimos, o "loop fonológico", é responsável pela retenção e manipulação da informação verbal; isto ocorre quando, por exemplo, alguém repete um número de telefone pra decorar antes de vertê-lo ao caderno. O outro, formado pelo esboço visual, é responsável armazenar e por manter as informações visuais (CARRUTHERS, 2015, p.75).

A localização dos sistemas ativados na memória de trabalho é um tópico controverso. Tradicionalmente, cita Carruthers, o componente executivo que controla e dirige os demais é localizado nos lóbulos frontais e, de início, também o eram o sistema fonológico e o esboço visual. Hoje, todavia, é sabido que a memória de trabalho possui formas para o cheiro, para o toque, a propriocepção etc., cada um deles implica recursos de áreas relevantes do cérebro. Mas um dos traços gerais observado em testes de imagem a partir de todas as formas de memória de trabalho foi a ativação distintiva de áreas sensoriais do cérebro. Os estudos de Kosslyn (CARRUTHERS, 2015, p.76) comprovam-no mediante uma estratégia diferente. Apoiado em uma extensa pesquisa empírica, Kosslyn mostra que os mesmos mecanismos de cima para baixo utilizados para a

percepção e para a classificação de um objeto são também empregados ao se imaginá-lo. As mesmas áreas de nível médio e alto na parte responsável pelo processamento visual do córtex temporal são ativadas nos dois tipos de atividade. Como essa imaginação ou fantasia faz parte das atividades do esboço visual da memória de trabalho, então a memória de trabalho, utilizando-se dos mesmos recursos, suporta as funções da imaginação e fantasia.

No capítulo 3.5, Carruthers explica o chamado conteúdo amodal envolvido na percepção. O argumento básico do autor consiste em extrair as implicações da refutação à distinção tradicional na filosofia entre percepção e cognição. Proponentes desta distinção acreditam que primeiro se percebe um objeto para depois se classificá-lo conceitualmente. Uma visão tradicional desse procedimento encontra-se, por exemplo, em Kant. Para Kant, a intuição sensível capta o múltiplo do fenômeno. É o entendimento, como operação fundamentalmente diversa, que reflete e abstrai os traços comuns dos objetos e depois redireciona a percepção sensível à conceituação deles. Neste sentido, vêem-se formas, sons, e figuras e cores e, só então, ato contínuo, atribui-se aos objetos uma identidade conceitual sob a forma de "isto é um livro vermelho etc." (ex).

Essa dicotomia tradicional contraria nossa intuição. Pois o que vemos, de fato, não são formas, cores e figuras isoladas e *depois* identificamos conceitualmente os objetos. Ao contrário, nós vemos figuras específicas - um carro vermelho, um livro de química, etc. No entanto, a mera intuição não tem força probante suficiente. Seria possível, afinal, pensar que ela está enganada. Que, ao sentirmos que já conceituamos aquilo que percebemos, estamos fundindo dois processos separados: a percepção sensorial e a conceituação, de tal modo que, se alguém anda em nossa frente, temos a percepção das formas e cores em movimento e a crença "alguém em minha frente está andando". Não é o caso. (CARRUTHERS, 2015, p.66) citando Kosslyn, menciona um modelo, ao qual ele atribui validade, que afirma processar-se um intercâmbio entre regiões visuais médias e regiões do córtex associacional, que estocam informações conceituais. Estes estudos demonstram refluírem informações conceituais para estágios de percepção, tornando-se componentes desses atos perceptivos. Uma percepção consciente é um híbrido conceitual-não-conceitual.

Mas não é só isso. As percepções visuais, processadas em áreas diferentes do cérebro, são organizadas de acordo com um modelo de arquivo de objetos (*object files*). Este modelo consiste em uma série de tipos crus de informações agrupadas da seguinte forma: imagine-se um tomate. As informações perceptivas agrupam-se: vermelho, redondo, macio, movendo-se para baixo. Se é correta a hipótese científica segundo a qual a percepção é um híbrido conceitual-não-conceitual, então as propriedades serão percebidas do seguinte modo: vermelho (vermelho conceito), redondo (conceito de redondo) etc. até que ocorra uma identificação - TOMATE CAINDO. Esta depende das partes - vermelho, redondo etc. - mas é algo acrescentado a elas. No exemplo seguinte, Carruthers menciona um dálmata. Vemos uma figura com várias manchas pretas; se desconhecemos o que elas desejam representar talvez haja um intervalo sensível até a percepção de que "é um dálmata!". Ora, mesmo enquanto essa identificação da unidade figurativa não se der, não deixaremos de ser conscientes das pintas, manchas, linhas com seus correspondentes conceitos, atrelados a elas em arquivos de objetos.

No que se refere à relação da memória com a consciência, estudos empíricos, afirma Carruthers, comprovam que a memória de trabalho é sensorialmente envolvente. Os seus conteúdos incluem imagens e informações conceituais; as primeiras, causadas por um fluxo de atenção dirigido aos conteúdos neurais e que aparece como atividade na área média sensorial, à semelhança do que ocorre na percepção ordinária.

Este conteúdo representacional está envolvido com o conteúdo sensitivo na memória de trabalho. Também se relacionam conteúdos representacionais e sensitivos nas atitudes proposicionais que, por sua vez, são co-dependentes da memória de trabalho igualmente. Carruthers dá o exemplo de uma praia: ao vermos uma bela praia tropical a concepção que dela formamos - composta de uma mistura inarredável de conteúdos representacionais e percepções não-conceituais - traz também imbricada certo *valor*. O mesmo ocorre quando deixamos livre a fantasia. Ao imaginar uma bela praia caribenha nós poderemos estar sentindo o desejo de ir até essa praia (o que, aliás, estou fazendo aqui a pensar em uma cerveja gelada).

As atitudes proposicionais não aparecem como destacadas e sobreimpostas à visão ou à imaginação, mas, sim, mescladas a elas. Do mesmo modo, ao fantasiarmos este objeto, podemos rapidamente passar para as considerações dos meios de obtê-lo; logo, o julgamento proposicional - eu decido parar de escrever e pegar a cerveja - está também mesclado às percepções sensoriais na memória de trabalho. O processo de alimentação e retroalimentação entre as áreas cerebrais responsáveis pela percepção e pela manutenção/manipulação do conteúdo da memória de trabalho explica porque essa interação ocorre.

Todos estes estados têm em comum, além do fato de serem compósitos de conteúdo conceitual e não-conceitual, a questão de serem conscientes. Em verdade, conforme anteriormente exposto, a relação entre memória de trabalho e consciência é estreita. Para autores como Carruthers e Baars, a memória de trabalho é um tipo de sistema de espaço global, mas esta classificação merece ser melhor qualificada.

A memória de trabalho não coincide com o espaço global (CARRUTHERS, 2015 p. 82-84). Algumas razões podem ser aventadas para sustentar essa tese. A mais imediata destas razões é o fato de a memória de trabalho e o espaço global não coincidirem conceptualmente. Esta razão, entretanto, não é suficiente; afinal, poderia ser o caso de expressões de significado diverso com a mesma referência. A posição citada de Prinz oferece uma justificativa a mais: se a memória de trabalho é ainda, essencialmente, memória, e tem como função sustentar os conteúdos, experiências efêmeras podem ocorrer sem que jamais tenham entrado na memória. Alguns estudos (Landman, 2003; Cowan, 2001) indicam ser a experiência consciente muito mais rica do que a memória de trabalho. A experiência consciente, aliás, pode ser de natureza fenomênica, levando à conclusão de que a relação entre fenomenalidade e memória de trabalho não implica uma identificação, nem mesmo ao nível da descrição científica de suas operações.

Estes aspectos mostram que, embora não se identifique com o espaço global, a memória de trabalho é notavelmente importante para a disponibilidade dos estados conscientes. Ela é uma das peças-chaves para se entender o fenômeno da consciência.

4 CONSIDERAÇÕES FINAIS

A discussão relativa às propriedades fenomênicas no capítulo anterior encerra o ponto principal de divergência entre Carruthers e Chalmers. Sem prejudicar o rumo da discussão, iremos analisar os argumentos desenvolvidos na última resposta que Carruthers fornece a Chalmers.

Em uma resposta ao artigo publicado mais tarde em *The Character of Mind*, Carruthers admite que a estratégia ali desenvolvida é bastante poderosa. De fato, a premissa segundo a qual P&~C ou é concebido ou não o é parece uma verdade necessária. Ele pondera ser uma condição para essa suposta necessidade que os conceitos fenomênicos (ou quase-fenomênicos) sejam tomados inequivocamente.

Um dos focos da réplica de Carruthers é a ideia de que os zumbis não compartilham a nossa situação epistêmica. Como vimos no capítulo anterior, Chalmers nega esse compartilhamento comparando valores de verdade e justificação das crenças dos zumbis e das cópias. Deste modo, um zumbi que possui o conceito fenomênico dado em C, ao conversar com outro dos seus, pode defender a tese de que a consciência não se esgota na descrição funcionalista. A proposição emitida pelo zumbi realista é falsa porque, no mundo de entes como ele, a consciência fenomênica de fato se reduz a uma manifestação da organização funcional. O seu outro exemplo apoia-se no experimento mental da Mary no quarto branco-e-preto. Se Mary zumbi sair do quarto, então ela possuirá um novo conhecimento, porém, este não será aquele de que “isto é vermelho”. Ela pode ter um conhecimento indexical – de que está, naquele momento e naquela circunstância – no estado psicológico correspondente à experiência do vermelho; pode ganhar a habilidade de classificar “vermelho” sob um conceito perceptual. Todavia, ela não obterá o conhecimento envolvido na própria experiência da cor (CHALMERS, 2010, p. 318).

Chalmers antecipa-se a um ataque dos que empregam a estratégia fenomênica. Concedendo algo ao defensor dessa posição, o autor imagina que os zumbis possam possuir outro estado em substituição aos estados fenomênicos, sobre os quais suas crenças tenham o mesmo valor de verdade

e sejam igualmente justificadas. Essas crenças são ditas “schmenomenais”, a respeito de estados de mesmo nome. Haveria uma analogia entre, de um lado, os zumbis schmenomenais e os fenomenais, e, de outro, entre os gêmeos das terras gêmeas onde água é H₂O e ZXY, respectivamente. Crenças sobre estados fenomenais seriam correspondentes a crenças sobre a água. Crenças sobre estados schmenomenais corresponderiam a crenças sobre a água gêmea. Como a função que estes estados ocupam na vida dos zumbis é a mesma desempenhada pelos fenomenais nas vidas das pessoas comuns, então ao obter conhecimento sobre uns e outros se obtêm a mesma quantidade de conhecimento, mas sobre objetos diferentes.

Carruthers pontua que Chalmers parece considerar pacífica a concepção externalista do significado. Com efeito, a interpretação externalista para o experimento das terras gêmeas é a seguinte: se Oscar e seu gêmeo têm experiências, respectivamente, com a água e a água gêmea, então o conteúdo de suas crenças nascidas desta experiência será diferente, mas os valores de verdade e demais aspectos epistêmicos relevantes serão os mesmos. Quando Oscar diz “A água é refrescante” e, com isso, refere-se a H₂O, e o Oscar gêmeo diz a mesma frase referindo-se a XYZ, ambas as crenças têm o mesmo valor de verdade neste e naquele mundo. Os conteúdos mudam, embora não seus valores de verdade. De igual maneira, quando Mary diz “Isto é vermelho”, ela utiliza um conceito fenomênico relativo a um estado fenomenal. Contudo, quando a Mary zumbi diz igual frase, ela utiliza um conceito schmenomenal, tocante a um estado schmenomenal.

Considerar que Mary zumbi e Mary real dizem a mesma coisa é uma confusão. Quando ambas afirmam “Isto é vermelho”, uma expressa uma experiência fenomenal e a outra um análogo disso – certa experiência schmenomenal, cuja função ocupa exatamente a mesma posição na economia epistêmica de Mary zumbi. Ambas possuem crenças verdadeiras, diferindo nos objetos discorridos.

Carruthers nota ser um modo de Chalmers tentar escapar a essa conclusão apelar para a intuição que temos sobre os zumbis. Estes não podem possuir nenhuma experiência introspectiva, em seu interior, tudo se encontra “escuro”. Mas, de fato, nada foi afirmado de contrário a essa pressuposição. Os

estados schmenomenais não são fenomenais, mas ocupam as mesmas posições epistêmicas; entretanto, não “iluminam” nada no foro interno do zumbi.

Em primeiro lugar, Chalmers poderia aprofundar essa argumentação tentando mostrar que, quando Mary e sua contraparte zumbi dizem, ambas, “Isto é a experiência de vermelho”, o que a primeira faz não é utilizar um conceito fenomenal conceitualmente isolado para capturar um estado.

A tese de Carruthers, apesar de interessante, possui uma falha fundamental. Com efeito, Chalmers argumenta (de forma relativamente curta e ineficaz) que a analogia entre estados schmenomenais e fenomenais não é boa, por conta de os primeiros não revelarem um âmbito interno rico em experiência, mas radicalmente pobre. O que não fica claro nessa argumentação é se haveria alguma implicação epistêmica decisiva a partir disso. Afinal, se o objetivo de Chalmers é provar que as situações epistêmicas de zumbis e humanos não são iguais, então, necessariamente, o ataque à analogia entre estados fenomenais e schmenomenais precisa incidir sobre a situação epistêmica, e não limitar-se a ponderar a relativa riqueza e pobreza experiencial dos estados schnomênicos e fenomênicos.

Acredito possuir alguns contra-exemplos bastante simples capazes de mostrar como não é possível construir todo tipo de proposição que faça menção a estados schmenomenais e fenomenais e tenha o mesmo valor de verdade em ambos os mundos. Mostrarei que, ao menos para algumas proposições, o valor de verdade muda de acordo com o mundo. Se verdadeiro no mundo normal, falso no mundo zumbi e vice-versa. Caso essa estratégia se revele frutífera, então estaremos atacando o coração do problema, o que a objeção de Chalmers terminou por elidir.

Suponha, por exemplo, que Chalmers e sua versão zumbi profiram a seguinte frase: “Isto é vermelho”. Neste caso, como mostrou Carruthers, as proposições enunciadas pelos dois agentes têm o mesmo valor de verdade. Ocorre apenas que “vermelho” para Chalmers refere-se a uma experiência fenomênica, ao passo que, para sua contraparte, ele remete a um estado shnomênico.

Não obstante, suponha que, desta vez, a frase seja “Ao ver que isto é vermelho, tenho uma experiência subjetiva”. Tal afirmação, proferida tanto por Chalmers quanto por sua versão zumbi, tem condições muito claras: se o agente profere a frase, e está em um estado descrito como sendo a posse de uma experiência subjetiva, então ela é verdadeira. Se não for o caso, é falsa.. A principal diferença entre elas é: no caso do zumbi, o “vermelho” visto refere-se a um estado schnomenal, e a posse de um conceito fenomênico, conceitualmente isolado de todos os outros, permite-lhe expressar-se desse modo. Já para Chalmers a experiência subjetiva de fato ocorreu. Por isso o valor de verdade das crenças parece não ser o mesmo.

Quando o zumbi-Chalmers reconhece o vermelho ele é capaz de emitir juízos sobre os objetos vermelhos, por exemplo. O zumbi-Chalmers pode julgar que certo objeto é vermelho; contudo, os estados schnomenais, por definição, não fornecem experiência subjetiva de nenhum tipo. Assim, a frase é necessariamente falsa no mundo alternativo dos zumbis. Em nosso mundo, porém, ela não só é verdadeira como, até, bastante óbvia: ao ver um objeto vermelho temos uma experiência que, corretamente, classificamos de subjetiva.

Uma resposta possível é tentar redefinir a experiência de acordo com a schnomenalidade. Neste mundo, a experiência subjetiva significaria algo distinto, significaria a capacidade de possuir um estado schnomenal e não um estado fenomenal. Dois são os problemas dessa saída.

O primeiro deles é que não fica claro por que a expressão ‘experiência subjetiva’ deveria depender da posse de um estado schnomenal. Imagine que os zumbis desejem falar sobre nós e imaginam não um zumbi-zumbi (como Carruthers propõe), mas, sim, alguém dotado da estranha propriedade (para eles) de possuir uma ‘experiência subjetiva’. Neste caso, eles utilizariam a mesma expressão com o mesmo significado, fazendo referência a um mesmo objeto. Assim, no uso comum da expressão para se reportarem a si mesmos, eles sabem que dizer “Ver que isto é vermelho é uma experiência subjetiva” é falso, já que sabem emprega-la em circunstâncias corretas justamente pelo fato de empregarem conceitos fenomenais para caracterizar

estados schnomenais, os quais, por definição, não podem fornecer experiência subjetiva.

O segundo problema é que, caso a experiência fosse assim definida, pareceria que o conceito de experiência depende do conceito puramente de reconhecimento que captura o estado schnomenal. Isto porque, em analogia com a experiência da terra gêmea, Oscar e o seu irmão empregam a mesma palavra com um significado distinto, determinado externamente pela referência. Conceitos que dependessem do de água poderiam ser então redefinidos de acordo com ele. Logo, a expressão 'água destilada' teria diferentes sentidos quando a água é H₂O e quando a água é XYZ. Assim, Ocorre, contudo, que essa analogia não é boa em se tratando de conceitos puramente de reconhecimento, afinal, se este tipo de conceito, por definição, está isolado de todos os demais, inclusive do de experiência (CARRUTHERS, 2000, p.59-64), então como tornar o conceito de experiência dependente de um conceito de puro reconhecimento? Esta redefinição não pode depender de nenhum conceito schnomenal. Assim, aparentemente, o zumbi-Chalmers está de posse apenas dos conceitos habituais para redefinir experiência. Mas, se é deles de que está de posse, por que haveria de redefini-la de forma estranha e aparentemente imotivada? Não o faria.

Outra objeção possível seria tentar repetir a distinção entre as abordagens de primeira e terceira pessoas. Deste modo, a frase "Ao ver que isto é vermelho tenho uma experiência subjetiva" confundiria as duas abordagens, recaindo, com isso, na crítica já feita. Muito embora não nos fique claro como haveria tal recaída, consideremos o ponto a fins de argumento. Ora, poderíamos reformular a frase para uma que indicasse uma terceira pessoa: "Ao ver que isto é vermelho *ele* tem uma experiência subjetiva". Poder-se-ia tentar escapar deste exemplo dizendo que não temos como afirmar tal frase com certeza, haja vista ser a experiência subjetiva de terceiros inacessível. Mas esse argumento não melhora a situação. Em nosso mundo, existe uma presunção muito forte e bastante razoável de que outros seres humanos *não* sejam zumbis. Se eu aponto alguém e profiro essa frase querendo indicá-lo como portador da experiência ela é, muito provavelmente, *verdadeira*. No mundo zumbi, inversamente, se os zumbis se julgassem possuidores de

experiência subjetiva eles incorreriam em um *erro maciço*; caso não julgassem, então a frase seria muito pouco provável. Seja a hipótese ou sua contraditória, o fato é que juízos a respeito da posse da experiência subjetiva parecem conduzir a uma assimetria entre o mundo zumbi e o nosso.

Outra possível objeção seria qualificar nossa estratégia de circular. Estaríamos, ao empregá-la, pressupondo aquilo que desejamos provar. Intentaríamos mostrar que os zumbis não podem proferir frases verdadeiras como “Esse copo de chá é feito de porcelana azul”, já que, desprovidos da experiência do azul, não poderiam afirmá-la. Carruthers objeta que eles podem, sendo sua crença igualmente verdadeira, e o conhecimento obtido, real, apenas tangente a *outro objeto*, qual seja, do estado *schnomenal* do azul. Assim, o mesmo processo se daria com o meu contra-exemplo. Eu estaria pressupondo aquilo que desejo provar: pressupondo uma assimetria, em que existem crenças com mesmo valor de verdade sobre *outro objeto*. Todavia, não estou fazendo exatamente aquilo que Carruthers ataca. Eu posso conceder que Carruthers esteja correto neste exemplo. Entretanto, e quando introduzo a noção de experiência subjetiva *na própria* frase? Se a crença do zumbi envolve assinalar a experiência subjetiva para si próprio ou para outrem, então essa crença é verdadeira somente se há uma experiência subjetiva.

Concedendo a noção de azul ser dada em termos puramente funcionais, como um estado *schnomenal* do ‘azul’ (do que tenho fortes dúvidas, mas aceito a introdução desse elemento pelo próprio Chalmers, novamente, a fins de argumento), não se pode dizer isso da experiência subjetiva. Uma experiência subjetiva dada em termos puramente funcionais seria, precisamente, o contrário de uma experiência subjetiva. Ainda se pode tentar contornar a crítica, dizendo que o zumbi Chalmers possuiria o conceito de experiência subjetiva sem que houvesse alguma. Ocorre que ele a assinala (e não a posse do seu conceito) quando assere “Ao ver que isto é vermelho tenho (ou ele tem) uma experiência subjetiva”.

Por outro lado, podemos proferir a frase “Ao ver o vermelho, tenho uma experiência subjetiva *schnomenal*”. Essa frase é falsa proferida na Terra, mas verdadeira proferida na terra-zumbi. E não há nenhuma dificuldade em se imaginar um contexto de asserção para ela. Do mesmo modo que nós

conseguimos identificar e classificar a nossa experiência subjetiva como sendo *fenomenal*, os zumbis poderiam classificar a deles como sendo *schmomenal*. Quais as implicações disso? Quando a minha crença envolve noção cuja classificação é crucial, então estar na terra ou estar na terra-zumbi é relevante para a determinação de seu valor de verdade.

Logo, se essa diferença é relevante para algumas crenças concebíveis, então existem diferenças relevantes nos valores de verdade de algumas crenças de homens e zumbis. E, se essas são de fato importantes, então a simetria epistêmica não se dá. Continua verdadeira a premissa do argumento de Chalmers, que prega não estarem os zumbis em nossa mesma posição epistêmica.

Comprova-se, ao final do trabalho, que: (1) Chalmers fornece um importante projeto de investigação acerca da ciência da consciência (2) Este projeto é ameaçado, em seu núcleo, por uma posição misterianista a respeito do problema. O misterianismo de McGinn tem uma característica singular em face das posições materialistas que Chalmers busca refutar. Ao contrário destas posições, ele recusa-se a compreender o problema por considerá-lo insolúvel. Desta maneira, julguei necessário enfrentar a posição misterianista de McGinn. (3) Ultrapassada esta etapa, apresento o sistema de Chalmers em seus traços mais salientes e típicos. É uma parte descritiva e exegética. (4) Apresento os resultados científicos de Carruthers, compatíveis metodologicamente com alguns pressupostos de Chalmers, notadamente, com o papel central da disponibilidade global dos estados conscientes e com a função que neles exerce a atenção (5) Conceitualmente, apresento as divergências entre Carruthers e o projeto anti-reducionista, com as quais Chalmers lidou ainda em um texto seu; sua reconstrução da estratégia dos conceitos fenomenais está exposta. (6) Apresento os resultados da investigação sobre a memória de trabalho, parte mais original e importante do livro *The Centered Mind* e, finalmente, nas considerações finais (7), encerro a dissertação com uma crítica original à réplica que Carruthers endereçou a Chalmers e que este último dispõe no seu site ainda sem resposta.

REFERÊNCIAS BIBLIOGRÁFICAS (TEXTOS CITADOS NESTE TRABALHO)

ARMSTRONG, D.M. (1968). *A Materialist Theory of the Mind* ed. Routledge.

BLOCK, Ned (1995). *On a confusion about a function of consciousness*. In: Behavioral and Brain Sciences, 18 (1995), pp. 227-287

BURGE, Tyler (2007). *Foundations of Mind*. New York: Oxford University Press

CARRUTHERS, Peter (2000). *Phenomenal Consciousness: a Naturalistic Theory*. Cambridge: Cambridge University Press

CARRUTHERS, Peter (2015). *The Centered Mind*. Oxford: Oxford University Press

CHALMERS, David (1996). *The Conscious Mind*. New York: Oxford University Press

CHALMERS, David (2010). *The Character of Consciousness*. New York: Oxford University Press

CHURCHLAND, Paul (1998). *Matéria e Consciência*. São Paulo: Ed. UNESP.

CHURCHLAND, Paul (2006). "Eliminative Materialism and the Propositional Attitudes". In: *The Journal of Philosophy*, Vol. 78, No. 2 (Feb., 1981), pp. 67-90

DAHEANE, Stanislas & CHANGEAUX, Jean-Pierre. "Experimental and Theoretical Approaches to Conscious Processing" In: *Neuron* 70, 2011, pp. 200-227

DENNETT, Daniel (1991). *Consciousness Explained*. New York: Bay Back Books, Little, Brown and Company

HARE, R.M. (1952). *The Language of Morals*. New York: Clarendon Press, Oxford

HORGAN, Terence (1993). "From Supervenience to Superdupervenience. Meeting the demands of a Material World". In: *Mind*, vol.102, 408, pp. 555-586

KIM, Jaegwon (1984). "Concepts of Supervenience". In: *Philosophy and Phenomenological Research* Vol. 45, No. 2, pp. 153-176

MCGINN, Colin (1987). "Can we Solve the Mind-Body Problem?" In: BLOCK, N. & FLANAGAN O. (eds) *The Nature of Consciousness*. Guven Guzeldere. MIT 1997

MCGINN, Colin (1997). *Problems in Philosophy. The Limits of Inquiry*. Cambridge, Massachusetts: Ed. Blackwell

MCGINN, Colin. (1996). *The Character of Mind. An Introduction to the Philosophy of Mind*. New York: Oxford University Press.

MACK & ROCK (1998). *Inattentional Blindness*.

MCGINN, Colin. (1999). *The Mysterious Flame*. New York: Basic Books

MORGAN, Loyd C. (1927). *Emergent Evolution*. London: Williams and Norgate.

NAGEL, Thomas (1974). "What is it like to be a Bat?" In: *The Philosophical Review* LXXXIII, 4, pp.435-450

PAPINEAU, David. (2008). "The Causal Closure of the Physical and Naturalism" In MCLAUGHIN, B. BECKERMANN, A., WALTER, S. (eds.) *The Oxford Handbook of Philosophy of Mind*. Oxford Press 2009 pp.53-65

PLACE, U.T, (1954). "Is Consciousness a Brain Process?" In: *Journal of British Psychology*, 47:1 pp.44-50

PUTNAM, Hilary (1967). "The Nature of Mental States". In LUDLOW, P. BEAKLEY, B. (eds.). *The Philosophy of Mind: Classical Problems, Contemporary Issue* Cambridge: The MIT Press pp. 97-105

REYNOLDS J. & CHELAZZI. L. "Attentional modulation of visual process" In: *Annual Reviews Neuroscience*. 27, pp.611-647

RYLE, Gilbert. (1959). *The Concept of Mind*. New York: ed. Routledge.

SEARLE, John. (1997). *The Mystery of Consciousness*. New York: The New York Review of Books

SERGENT, Claire & DEHAENE, Stanislas (2004). "Is Consciousness a Gradual Phenomenon? Evidence for an All-or-None Bifurcation During the Attentional Blink". In: *Psychological Science*, vol.15, nº11, pp.720-728

SMART, J.J.C. (1959) "Sensations and Brain Processes". In: *British Journal of Psychology* XLVII. pp.141-156

SPERRING, Paul. (2004). "Should We Give Up on Reductive Physicalism?" In: *Richmond Journal of Philosophy* 8, pp.1-7

TEIXEIRA. (1997). "A teoria da Consciência em David Chalmers" In: *Psicologia USP* vol. 8 n. São Paulo