



**UNIVERSIDADE FEDERAL DE ALAGOAS**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
Coordenadoria de pós-graduação  
Programa de Pós-graduação em Filosofia

**A TESE FUNCIONALISTA DE DANIEL DENNETT**  
**A mente como um software rodando no cérebro**

**JOSÉ ALFREDO MELO DOS SANTOS**

29 de novembro de 2023  
Maceió – Al

**JOSÉ ALFREDO MELO DOS SANTOS**

**A TESE FUNCIONALISTA DE DANIEL DENNETT**  
**A mente como um software rodando no cérebro**

**Dissertação de Mestrado** apresentada ao Programa de Pós-Graduação em Filosofia da UFAL (área de concentração: Linguagem e Cognição), como parte dos requisitos necessários para a obtenção do Título de Mestre em Filosofia.



Documento assinado digitalmente  
RICARDO SEARA RABENSCHLAG  
Data: 10/01/2024 16:14:42-0300  
Verifique em <https://validar.itl.gov.br>

---

Prof. Dr. Ricardo Seara Rabenschlag  
Orientador

Maceió – Al, 29 de novembro de 2023

**JOSÉ ALFREDO MELO DOS SANTOS**

**A TESE FUNCIONALISTA DE DANIEL DENNETT**

**A mente como um software rodando no cérebro**

**Dissertação de Mestrado** apresentada ao Programa de Pós-Graduação em Filosofia da UFAL (área de concentração: Linguagem e Cognição), como parte dos requisitos necessários para a obtenção do Título de Mestre em Filosofia.

Trabalho aprovado. Maceió – AL, 29 de novembro de 2023:

Documento assinado digitalmente  
 **RICARDO SEARA RABENSCHLAG**  
Data: 10/01/2024 16:14:42-0300  
Verifique em <https://validar.iti.gov.br>

---

**Prof. Dr. Ricardo Seara Rabenschlag**

**Orientador**

Documento assinado digitalmente  
 **ANDRE LUIZ DE ALMEIDA LISBOA NEIVA**  
Data: 10/01/2024 18:20:01-0300  
Verifique em <https://validar.iti.gov.br>

---

**Prof. Dr. André Luiz A. Lisboa Neiva**

**Convidado 1**

Documento assinado digitalmente  
 **TARIK DE ATHAYDE PRATA**  
Data: 12/01/2024 20:16:36-0300  
Verifique em <https://validar.iti.gov.br>

---

**Prof. Dr. Tárík de Athayde Prata**

**Convidado 2**

Maceió – Al  
2023

**Catlogação na fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**  
Bibliotecária: Taciana Sousa dos Santos – CRB-4 – 2062

S23t Santos, José Alfred Melo dos.  
A tese funcionalista de Daniel Dennett: a mente como um software rodando no cérebro / José Alfredo Melo dos Santos. – 2023.  
130 f.: il.

Orientador: Ricardo Seara Rabenschlag.  
Dissertação (Mestrado em Filosofia) – Universidade Federal de Alagoas. Instituto de Ciências Humanas, Comunicação e Artes. Programa de Pós-graduação em Filosofia. Maceió, 2023.

Bibliografia: f. **128-130**.

1.Relação mente-cérebro. 2 Dennett, Daniel Clement, 1942-. 3. Consciência. I. Título.

CDU: 13:159.922

“A consciência deve ser reconhecida como um aspecto conceitualmente irreduzível da realidade que está necessariamente conectado a outros aspectos igualmente irreduzíveis – como os campos eletromagnéticos são irreduzíveis, mas necessariamente conectados ao comportamento de partículas carregadas e campos gravitacionais com comportamento de massas, e vice-versa.”

Thomas Nagel

“A curiosidade do ser humano e a perspicácia de sua razão revelaram boa parte do que a natureza sempre manteve oculta. A estrutura do espaço-tempo, a constituição da matéria, as muitas formas de energia, a natureza da própria vida – todos esses mistérios se tornaram livros abertos para nós. Sem dúvida, existem questões profundas que continuam sem respostas e revoluções que ainda estão à nossa espera, mas seria difícil exagerar a explosão de conhecimentos científicos que os seres humanos produziram nos últimos quinhentos anos.

Apesar do avanço geral, um mistério fundamental permanece, em grande parte, ainda um mistério: a natureza da inteligência consciente.”

Paul M. Churchland

“No futuro distante vejo áreas abertas a pesquisas muito mais importantes. A psicologia terá base em um novo alicerce, o da aquisição necessária de cada poder mental e capacidade por transições graduais. A origem do homem e sua história serão esclarecidas.”

Charles Darwin

“Não há nenhuma célula ou grupo de células no cérebro com tal preeminência anatômica ou funcional que pareça ser a pedra angular ou o centro de gravidade de todo o sistema.”

William James

“Observo algo e procuro uma razão: originalmente isto significa que procuro uma intenção nele, e acima de tudo alguém que tenha intenções; um sujeito, um agente para cada evento, uma ação – anteriormente víamos intenções em todos os eventos, este é o nosso hábito mais antigo. Os animais também o possuem?”

Friedrich Nietzsche

## RESUMO

Este projeto tem por finalidade examinar o problema da relação mente-cérebro – uma querela ontológica e milenar examinada hoje não só pela filosofia, mas também por outras áreas de pesquisas, tais como a neurociência, a ciência cognitiva e a inteligência artificial – sob a ótica do filósofo estadunidense Daniel C. Dennett. Tendo como embasamento o seu texto *Consciousness Explained*, de 1991, iremos seguir a trajetória que o levou a considerar a consciência com um ninho de memes, um pandemonium de ficções úteis. Veremos como, em seu pensamento, procura subjugar o dualismo cartesiano (e todo seu legado) nos apresentando um fisicalismo de dupla face, refletido na relação *software-hardware*. Inicialmente, iremos abordar algumas interfaces que servem de valimentos para fundamentar o funcionalismo homuncular dennettiano. Em virtude disso, iremos realizar um breve exame sobre as pesquisas realizadas nas ciências cognitivas, na neurociência e na tese darwiniana sobre a evolução das espécies, bem como verificar como Dennett realiza um aproveitamento das descobertas nessas áreas. Além disso, analisaremos um recurso amplamente utilizado pelos filósofos na edificação de teorias – os chamados experimentos de pensamento. Na sequência, abordaremos três dos principais conceitos dennettianos na edificação de sua teoria da consciência, a saber: o método heterofenomenológico, o meme e os sistemas intencionais. Por fim, os reacionários à teoria de Dennett, segundo os quais destacamos John Searle e David Chalmers, serão objeto de estudo. Concluiremos nossa exposição com um exame da teoria da identidade de Dennett, o *self*. Ao que parece, o pensamento de Dan Dennett é um bom recurso introdutório para qualquer pesquisador que queira se aventurar nesse tema tão complexo que é a natureza da inteligência consciente.

**Palavras-chave:** Cérebro. Dennett. Meme. Consciência. Mente.

## **ABSTRACT**

This project aims to examine the problem of the mind-brain relationship – an ontological and ancient dispute examined today not only by philosophy, but also by other areas of research, such as neuroscience, cognitive science and artificial intelligence – from the perspective of the American philosopher Daniel C. Dennett. Based on his 1991 text *Consciousness Explained*, we will follow the trajectory that led him to consider consciousness as a nest of memes, a pandemonium of useful fictions. We will see how, in his thinking, he seeks to subjugate the Cartesian dualism (and all its legacy) by presenting us with a double-sided physicalism, reflected in the software-hardware relationship. Initially, we will address some interfaces that serve as validities to support Dennett's homuncular functionalism. As a result, we will carry out a brief examination of research carried out in cognitive sciences, neuroscience and Darwin's thesis on the evolution of species, as well as verify how Dennett makes use of discoveries in these areas. In addition, we will analyze a resource widely used by philosophers in theory building – the so-called thought experiments. Next, we will address three of Dennett's main concepts in the construction of his theory of consciousness, namely: the heterophenomenological method; the meme; and intentional systems. Finally, the reactionaries to Dennett's theory, according to which we highlight John Searle and David Chalmers, will be the object of study. We will conclude our exposition with an examination of Dennett's theory of identity, the self. Apparently, Dan Dennett's thinking is a good introductory resource for any researcher who wants to venture into this very complex topic that is the nature of conscious intelligence.

**Keywords:** Brain. Dennett. Meme. Conscience. Mind.

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	09
<b>1. OS VALIMENTOS DENNETTIANOS</b> .....	15
1.1 As estranhas inversões: Darwin e Turing .....	22
1.2 O papel da ciência cognitiva e da I.A. ....	29
1.3 A neurociência e o problema fácil .....	36
1.4 Experimentos de pensamento: bombas de intuições ou muletas de lança? .....	47
<b>2. UMA TEORIA EMPÍRICA DA CONSCIÊNCIA</b> .....	57
2.1 Crítica ao dualismo da substância e ao fisicalismo ingênuo .....	60
2.2 O método heterofenomenológico .....	65
2.3 A tese memética: a consciência como um pandemonium .....	68
2.4 A formulação da teoria dos sistemas intencionais .....	81
<b>3. REAÇÃO AO FUNCIONALISMO HOMUNCULAR</b> .....	87
3.1 O embate com Chalmers: será que há o problema difícil? .....	90
3.2 "Cerceando" Searle .....	99
3.3 Self: o centro de gravidade narrativa .....	108
<b>CONSIDERAÇÕES FINAIS</b> .....	120
<b>BIBLIOGRAFIA</b> .....	128

## INTRODUÇÃO

O cerne de nossa investigação será a natureza da inteligência consciente, sob a ótica de Daniel Clement Dennett III, filósofo estadunidense contemporâneo, haja vista que se mostra como um bom ponto de partida para qualquer pesquisador que queira se aventurar nesse tema tão complexo, ou seja, para aquele que se propõe a resolver, ou pelo menos se debruçar, sobre o problema da relação mente-cérebro.

No início de *Consciousness Explained*, de 1991, é delineada a metodologia para descrever a consciência. De forma que no “experimento de pensamento sobre o cérebro na cuba e na analogia com o jogo da psicanálise são exercícios de aquecimento para a tarefa principal, que é esboçar uma teoria dos mecanismos biológicos” (Dennett, 1991a, p. 17 – nossa tradução). Esse intento é calcado em uma abordagem epistemológica interdisciplinar.

Para realizar tal objetivo buscaremos, a princípio, examinar a crítica realizada por Dennett à abordagem da corrente dualista da mente. O filósofo nutre uma aversão a essa abordagem por ela, no seu entendimento, admitir que mente e cérebro são duas coisas distintas, mas que se relacionam de alguma maneira misteriosa. Nesse sentido, apresentaremos a tese do dualismo cartesiano e os problemas inerentes a esta corrente filosófica apontada pelo filósofo.

Em seguida, iremos analisar alguns textos de Dennett, a fim de delinear a estrutura e as motivações de sua tese, que caracterizam seu pensamento como *funcionalismo homuncular*<sup>1</sup>, mas que, ao mesmo tempo, estabelece uma crítica ao *funcionalismo mecânico*<sup>2</sup> de Jerry Fodor. Nesse sentido, para Dennett, a mente (que é do nível do virtual) admite vários tipos de dispositivos materiais e que, no entanto, não é caracterizado por nenhum deles.

Buscaremos deixar claro, também, a relevância de uma teoria da linguagem no pensamento dennettiano com a discussão sobre sua *teoria memética*, bem como, sua articulação com o conceito de *sistema intencional* – termos basilares na teoria de Dennett. É aqui que reside a hipótese de que a linguagem – caracterizada por uma relação de transmissão e recepção de termos e dados culturais – depende de uma organização de matéria específica que promova

---

<sup>1</sup> Uma das várias abordagens do problema mente/cérebro. Pressupõe que os estados mentais se estabelecem por uma definição e caracterização segundo o papel funcional que exercem no trajeto que envolvem o *input* e o *output* – a caixa preta – de um organismo ou sistema, ou seja, é a sua atribuição na economia comportamental do sujeito (cf. Tripicchio, 2004, p. 86). A rigor, trata-se de uma “concepção que entende que os dados comportamentais devem ser vistos como manifestações de aptidões psicológicas do sujeito, e essas aptidões devem ser explicadas pela compreensão do sujeito como sistema de componentes interligados” (Bunnin; Tsui-James, 2002, p. 176). Em Dennett, o funcionalismo ganha contornos não-reducionista.

<sup>2</sup> O funcionalismo mecânico, assim denominado por Bunnin e Tsui-James (2002), pode ser resumido como aquele que concebe a mente como algo que pode ser comportado por vários tipos de dispositivos materiais que, no entanto, não é caracterizado por nenhum deles.

uma resposta comportamental consciente, ou seja, um arranjo de matéria que possa se organizar em uma estrutura que esteja apta a promover *inputs* e *outputs*, sem que isso dependa de outra consciência.

Assim, ao investigarmos as reflexões de Dennett sobre a natureza da mente, necessariamente entraremos com mais vigor na seara das pesquisas científicas vinculadas aos estudos da ciência computacional, da biologia darwiniana, da ciência cognitiva e da neurociência. Ele aponta em *Brainstorms*, de 2006, seu distanciamento do behaviorismo e uma aproximação das ciências cognitivas, pois, “o bom senso diz que o behaviorismo está morto e a ciência cognitiva, uma aliança de psicologia cognitiva, linguística e inteligência artificial, é a onda do futuro” (Dennett, 2006, p. 25-26). Sobre o behaviorismo Dennett (1991a, p. 183 – nossa tradução) aduz que:

por muitos anos no século XX, a teoria mais influente foi o behaviorismo de B. F. Skinner, no qual os pares estímulo-resposta eram os candidatos à seleção, e os estímulos "reforçadores" eram o mecanismo de seleção. O papel dos estímulos prazerosos e dolorosos – a cenoura e a vara – na modelagem do comportamento é inegável, mas o mecanismo de "condicionamento operante" do behaviorismo foi amplamente reconhecido como muito simples para explicar as complexidades da fixação de design pós-natal em espécies tão complicadas quanto os seres humanos (e provavelmente em pombos também, mas isso é outra questão).

No que tange às falhas do cartesianismo (o teatro cartesiano), a teoria da consciência dennettiana surge como alternativa a ser considerada. Seu modelo exige que abandonemos a nossa zona de conforto para abraçar uma reflexão radicalizada da ideia comum de fluxo de consciência, um modelo que é, a princípio, amplamente de caráter contra-intuitivo, mas que se torna cada vez mais palatável, à medida que somos instruídos a partir dos estudos sobre o funcionamento do cérebro “que foram ignorados até agora pelos filósofos – e cientistas” (Dennett, 1991a, p. 17 – nossa tradução). Quanto às características da sua metodologia, Dennett (1991a, p. 40 – grifo nosso) a descreve da seguinte forma:

(1) Não é permitido o tecido *Wonder*. Tentarei explicar cada aspecto intrigante da consciência humana dentro da estrutura da ciência física contemporânea; em nenhum momento farei um apelo a forças, a substâncias ou a poderes orgânicos inexplicáveis ou desconhecidos. Em outras palavras, pretendo ver o que pode ser feito dentro dos limites conservadores da ciência padrão, salvando um apelo por uma revolução no materialismo como último recurso.

(2) Sem anestesia simulada. Já foi dito dos behavioristas que eles fingem anestesia – eles fingem que não têm as experiências que sabemos muito bem que eles compartilham conosco. Se eu quiser negar a existência de algum aspecto controverso da consciência, o ônus recai sobre mim para mostrar que é de alguma forma ilusória.

(3) Sem picuinhas sobre detalhes empíricos. Tentarei acertar todos os fatos científicos, na medida em que são conhecidos hoje, mas há muita controvérsia sobre quais avanços excitantes resistirão ao teste do tempo. Se eu me restringisse a “fatos que chegaram aos livros didáticos”, não poderia me valer de algumas das descobertas recentes mais reveladoras (se é isso que são). E eu ainda acabaria fornecendo

involuntariamente algumas falsidades, se a história recente servir de guia. Algumas das “descobertas” sobre a visão pelas quais David Hubel e Torstein Wiesel receberam merecidamente o Prêmio Nobel em 1981 estão agora sendo desvendadas, e a famosa teoria “retinex” da visão de cores de Edwin Land, que tem sido considerada pela maioria dos filósofos da mente e outros não especialistas como fato estabelecido por mais de vinte anos, não é tão bem visto entre os cientistas visuais.

A esta altura o leitor já percebeu que o método aplicado, na teoria dennettiana, será a *Abordagem-de-cima-para-baixo*<sup>3</sup>. O problema da mente é analisado a partir da manifestação do comportamento, ou seja, o filósofo examina a questão sobre a natureza da inteligência consciente a partir da “compreensão atual do que as criaturas inteligentes fazem e depois se pergunta que tipo de operações subjacentes poderiam produzir ou explicar essas atividades cognitivas” (Churchland, 2004, p. 156-57). Contudo, sua aproximação ao behaviorismo, apesar de notória, é parcial e esperamos esclarecer esse ponto no decorrer desse trabalho. Tentamos condensar o pensamento de Dennett em três capítulos que, acredito, estabelece uma mediação com a obra desse pensador, concernente a uma teoria da consciência.

*Os valimentos dennettianos* é o tema do capítulo I. Nesse capítulo, trataremos do exame de algumas ideias – desenvolvidas nas ciências empíricas e na filosofia – que Dennett recorre para fundamentar sua tese da consciência. Seu pensamento é amplamente estruturado considerando as “estranhas inversões” realizadas por Charles Darwin (1809-1882) e Alan Turing (1912-1954). De Darwin, foi explorado o conceito de evolução natural – esse conceito é assimilado como uma ideia tão genial que pode ser aplicado em várias áreas de pesquisas, inclusive em uma teoria da mente. De Turing, é extraída a concepção segundo a qual é possível realizar a produção de inteligência a partir de um sistema composto por partes ignorantes, observando um processo algorítmico. O papel da ciência cognitiva e da inteligência artificial (I.A.) também será objeto de investigação. Examinaremos como as ciências cognitivas, apesar de não serem aqui o ponto mais importante dessa análise, se aproximam da concepção reducionista. O núcleo de suas investigações reside na tentativa de entender como sistemas artificiais podem ou não produzir estados conscientes. Nesse sentido, os trabalhos realizados nessas áreas são canalizados para o desenvolvimento de simulações de atividades mentais, características dos seres humanos. Trata-se de uma ciência que auxilia a I.A. no desenvolvimento de processos de aprendizagem artificial. No ponto concernente à inteligência artificial, iremos examinar como Dennett utiliza os experimentos realizados nessa área –

---

<sup>3</sup> O método de *Abordagem-de-baixo-para-cima* analisa as estruturas do sistema nervoso que podem produzir a atividade mental. Esse método é mais próximo da neurociência (cf. Churchland, 2004, p. 156).

desenvolvidos a partir de 1950 – para fundamentar o critério funcional como limitador do âmbito mental.

A neurociência não ficará de fora da nossa análise. Ela será examinada sob dois aspectos: primeiro queremos verificar quais são as contribuições dos estudos dessa área para uma teoria da mente e, em seguida, demonstrar que os denominados “problemas fáceis”<sup>4</sup> – como, por exemplo, tentar entender se a função da memória é uma instância centralizada ou distribuída por todas regiões do cérebro e tentar entender o que ocorre no cérebro quando está sob efeito de alucinógenos –, apontados por David Chalmers, na verdade, não são nada fáceis.

Ainda nesse capítulo, nos atrevemos a entrar na seara das discussões sobre os experimentos de pensamento. Nessa seção temos por desígnio examinar o modo pelo qual os experimentos de pensamento, enquanto metodologia ontológica e/ou epistemológica, contribuem para uma ontologia da mente. É por esse tipo de recurso que Dennett elabora as réplicas a seus críticos, bem como constrói argumentos para defender sua tese. Observamos uma leve apologia realizada pelo filósofo, segundo a qual os experimentos de pensamento são como bombas de intuição. Nessa perspectiva, os experimentos de pensamento<sup>5</sup> ocupam um papel relevante em uma teoria da mente.

No capítulo II, sob o título de *uma teoria empírica da consciência*, procuramos retomar, inicialmente, a crítica de Dennett ao dualismo cartesiano, visando preparar o terreno para os conceitos que estruturam sua tese, que servem de pilares para sua teoria da consciência. Em detrimento à “contemplação” promovida pelo teatro cartesiano, o *método heterofenomenológico* emerge como uma alternativa interessante. Conquanto a notória similaridade com o behaviorismo seja patente, acreditamos conseguir demonstrar que há diferenças importantes. A *teoria memética* será examinada logo em seguida. Segundo o funcionalismo homuncular dennettiano, é ela que proporciona aquilo que é interpretado por muitos como sendo os conteúdos mentais. A rigor, ela se configura como a teoria da linguagem de Dennett. O meme é sem dúvida um dos pilares do pensamento dennettiano e assume o rótulo de conceito mais controverso de sua teoria. Encerramos o capítulo examinando a formulação da teoria dos *sistemas intencionais*, segundo a qual é possível atribuir um comportamento

---

<sup>4</sup> Chalmers (1996) usa essa terminologia para realizar uma comparação com aquilo que entende ser extremamente complicado de explicar, o denominado problema difícil da consciência.

<sup>5</sup> Considerando a taxonomia realizada por Robert Brown (1991), acreditamos ser possível delinear a importância desse tipo de experimento. Essa classificação descreve dois tipos gerais de experimentos de pensamento, a saber: construtivos e destrutivos. O primeiro remete a um intento de promover uma refutação de uma teoria. Já o segundo, vislumbra corroborar uma tese facilitando seu entendimento. Contudo, não deixaremos de mencionar as críticas a essa taxonomia.

inteligente a partir da análise dos três tipos de posturas das entidades, a saber: postura física, postura de design, e postura intencional.

Finalmente, no capítulo III, uma vez estabelecido um entendimento das características gerais dos experimentos de pensamento, examinaremos as críticas direcionadas ao pensamento de Dennett, que configuram uma *reação*<sup>6</sup> ao *funcionalismo homuncular*. Considerando alguns experimentos de pensamento clássicos (como o argumento do mundo zumbi de David Chalmers e o argumento do quarto chinês de Searle), iremos analisar as críticas direcionadas à algumas teses basilares dennettiana, como por exemplo o aspecto ontológico da tese memética. Essa saga se inicia com uma discussão sobre o problema difícil delineado por Chalmers e a crítica ao conceito de zumbi. Com o objetivo de refutar qualquer tipo de fisicalismo, o filósofo australiano elabora um experimento mental calcado na ideia da possibilidade de se imaginar um mundo fisicamente igual a nosso em todos os aspectos – inclusive no que tange a estrutura cerebral –, mas que, apesar da anatomia cerebral ser idêntica à nossa, não produz uma consciência.

O naturalismo biológico de John Searle (1998) também terá como pano de fundo uma crítica ao funcionalismo dennettiano. Para tentar refutar a analogia entre o cérebro e um computador (tese da *inteligência artificial forte*), realizada por Dennett, ele sugere também um experimento mental. Trata-se do famoso argumento do quarto chinês. Nesse experimento, Searle tenta demonstrar que o processamento de dados do computador necessita apenas de uma estrutura sintática para seu funcionamento, enquanto que, nos cérebros humanos, há a necessidade de uma estrutura linguística que envolve semântica, sintaxe e pragmática.

Encerramos nossa análise investigando o *self* dennettiano. A concepção de consciência dennettiana tem como uma de suas implicações a necessidade de explicar o que seria o “eu”. A edificação de uma teoria da identidade pessoal aponta para a ideia de que o *self* é uma espécie de centro de gravidade narrativa. Veremos que há um problema nessa proposta, pois, não fica claro qual seria a natureza desse conceito.

A filosofia da mente é um dos grandes campos de discussão da atualidade – tanto da filosofia quanto da ciência. Talvez a mente humana seja apenas uma dentre tantas outras existentes. Talvez devêssemos considerar a possibilidade de que existam outros seres que possuem uma mente, em níveis diversos (quem sabe as plantas e outros animais) como os bebês,

---

<sup>6</sup> Dennett (1999) chama atenção para uma gama de opositores ao naturalismo evolucionista e mecanicista, que emergiu nas últimas décadas do século passado – nos limitaremos a examinar os argumentos de David Chalmers e John Searle.

por exemplo, que não possuem uma linguagem bem articulada. A esse respeito Dennett (1997, p. 16) assevera:

Nós, seres humanos, compartilhamos um mundo subjetivo – e sabemos disso – de uma maneira que está inteiramente além das habilidades de quaisquer outras criaturas no planeta, pois sabemos falar uns com os outros. Os seres humanos que não tem (ainda) uma língua com a qual comunicar-se são a exceção, e esta é a razão pela qual temos um problema particular ao tentar entender como é ser um bebe recém-nascido ou um surdo-mudo.

Pode-se perceber que as questões que envolvem a relação mente-cérebro se enquadram como algumas das mais complexas. O labor aqui empregado tem por objetivo ajudar a compreender como a abordagem de Daniel Dennett se propôs a resolver o problema ontológico<sup>7</sup> referente à mente e como sua teoria dos memes se relaciona com uma teoria neurocientífica. Entendemos que examinando os conceitos centrais da teoria do autor daremos margem para uma melhor compreensão do problema mente-cérebro, ressaltando assim a importante contribuição desse pensador. Com isso, seremos capazes de desenvolver um pensamento reflexivo mais apurado acerca de questões que envolvem o estudo da mente. Dennett, (1991a, p. 23 – nossa tradução) trata a questão ainda mais a fundo:

Embora ainda existam pensadores que defendem corajosamente que a consciência seja uma coisa genuína e preciosa (como o amor, como o ouro), uma coisa que é apenas “óbvia” e muito, muito especial, cresce a suspeita de que isso seja uma ilusão. Talvez os vários fenômenos que conspiram para criar o sentido de um único fenômeno misterioso não tenham unidade mais fundamental ou essencial do que os vários fenômenos que contribuem para a sensação de que o amor é uma coisa simples.

Nesse sentido, acreditamos ser importantíssimo iniciarmos nossa investigação a partir da consideração dos vários níveis hierárquicos da reflexão dennettiana, calcada em uma ampla cadeia de citações que, ao nosso entender, esclarece o aspecto geral segundo o qual foi edificado o projeto filosófico de Dennett. Uma abordagem que não estabelece uma linha divisória entre filosofia e ciência.

Portanto, a nossa intenção aqui não será a de estabelecer uma resposta definitiva para o problema ontológico da consciência, nem realizar uma apologia ou uma crítica fundamentada em negações de fatos esclarecido pelo pensamento dennettiano, mas, sim, de realizar uma leitura significativa que permita um bom entendimento do problema, ou seja, que facilite a visão do horizonte das pesquisas sobre esse tema tão complexo que é a relação mente-cérebro. Não pretendemos gerar novas polêmicas, porém, não vamos esconder novos conflitos.

---

<sup>7</sup> Considera-se que “uma questão ontológica é simplesmente uma questão sobre que coisas realmente existem e qual sua natureza essencial” (Churchland, 2004, p. 17).

## 1. OS VALIMENTOS DENNETTIANOS

Neste capítulo, iremos expor algumas das disciplinas que Dennett recorre para fundamentar sua tese sobre a natureza da inteligência consciente. Veremos como as pesquisas de algumas das principais interfaces são relevantes para sua teoria da mente e como o filósofo insere algumas das descobertas nessas áreas em sua pesquisa.

Dentro da perspectiva que distingue cérebro e mente, algumas perguntas se fazem pertinentes: como podemos explicar a relação entre um objeto físico (o cérebro) e algo não-físico (a mente)? Até que ponto existe uma hierarquia nessa relação? Se mente e cérebro são coisas diferentes, qual seria o ponto que estabelece essa ligação? Por que danos cerebrais influenciam o comportamento da mente? As tentativas de respostas a essas questões nos são dadas, a princípio, por duas correntes que fazem abordagens fundamentalmente opostas, a saber: o dualismo e o materialismo – também denominado de monismo<sup>8</sup>. Nos debates atuais sobre a mente, o termo fisicalismo é comumente utilizado para expressar uma noção mais ampla de materialismo<sup>9</sup>, ou seja, os teóricos da mente, cada um ao seu modo, concebem o fenômeno mental como parte integrante do mundo físico. Nesse sentido, os próprios fenômenos psicológicos são considerados como fenômenos físicos.

O problema ontológico que permeia a relação entre a mente e o cérebro vem intrigando os estudiosos do tema há séculos. Dentre as grandes contribuições oferecidas para essa discussão, encontramos nomes como o de René Descartes (1596-1650) e sua teoria da substância pensante, considerado por muitos o precursor da filosofia da mente e um dos principais nomes da corrente dualista<sup>10</sup>. Já entre os grandes nomes do debate contemporâneo<sup>11</sup>, encontramos, além de Dennett, John Searle e seu naturalismo biológico e o dualismo naturalista de David Chalmers.

---

<sup>8</sup> Os monistas são os filósofos que admitem apenas a existência de apenas um tipo de entidade, ou seja, podem ser tanto os materialistas quanto os idealistas. No entanto, o termo foi monopolizado pelos teóricos da mente de tendência materialista (cf. Abbagnano, 2012, p. 794).

<sup>9</sup> Embora não seja rotulada como tal, a teoria searleana é um ótimo exemplo de uma postura materialista que procura ampliar a concepção de fisicalismo.

<sup>10</sup> Dennett nutre grande rejeição a essa corrente filosófica. Chega a defender que o dualismo (o ponto de vista de que as mentes são compostas de uma substância não física e absolutamente misteriosa) e o vitalismo (o ponto de vista de que as coisas vivas contêm algum tipo de substância física, mas igualmente misteriosa – o elã vital) foram, ou pelo menos deveriam ser, relegados ao monte de lixo da história, junto com a alquimia e a astrologia (cf. Dennett, 1997, p. 22).

<sup>11</sup> Podem ser incorporados a essa lista nomes como o de Thomas Nagel, Terence Horgan e Paul Churchland.

Podemos nos apropriar de uma visão mais panorâmica da complexidade da análise da ontologia da inteligência consciente, com o auxílio de Paul Churchland em sua obra *Matéria e Consciência*, de 2004.

Segundo Churchland (2004), a primeira abordagem moderna, realizada no que tange o problema da relação da alma<sup>12</sup> com o corpo, é creditada ao filósofo francês René Descartes e seu *dualismo da substância*. Ao estabelecer a *dúvida hiperbólica* como método para alcançar sua primeira certeza – o *cogito* cartesiano –, Descartes torna-se o precursor do racionalismo moderno. Assim, claro está, os pensadores seiscentistas já se debruçaram sobre o problema de como a alma pode interagir causalmente com o corpo.

Conforme Dennett (1991a), Descartes concebe o homem como um ser constituído por duas substâncias diversas e hierárquicas: a coisa pensante, a *res cogitans*, que seria a natureza essencial da inteligência, sua caracterização se dá essencialmente pelo pensamento; e a coisa extensa, a *res extensa*, que seria a matéria, que no cartesianismo é caracterizada não tanto por sua materialidade, mas, sobretudo, por sua extensão – os argumentos da sexta meditação (argumento da divisibilidade e argumento metafísico) mostram que o pensamento exclui a extensão. Em termos concretos, isso significa que tais substâncias, estranhamente, se influenciam mutuamente.

Seguindo a trilha de análise de Churchland (2004), examinemos brevemente algumas variantes do dualismo da substância, como o: *dualismo popular* – que é uma versão mais branda do dualismo da substância, caracterizada pela concepção de que há um fantasma que controla o corpo; e o dualismo da propriedade – cujo defensores afirmam que o cérebro é dotado de propriedades especiais que são de caráter não-físico. O Epifenomenalismo é uma variante desse tipo de dualismo. Sobre essa concepção de consciência, Churchland (2004, p. 31) nos fornece mais detalhes:

Os fenômenos mentais não constituem uma parte dos fenômenos físicos no cérebro determinando, em última análise, nossas ações e comportamentos, mas, ao contrário, eles ocorrem, por assim dizer, “acima do embate”. Os fenômenos mentais são, dessa forma, epifenômenos. Sustenta-se que eles se manifestariam ou viriam à tona apenas a partir do momento em que o cérebro em desenvolvimento ultrapassa um certo nível de complexidade.

A mente como epifenômeno da atividade cerebral não produz nenhum efeito causal. Esse efeito emana exclusivamente das atividades físicas no cérebro. O fato de nossas volições coincidirem com essas atividades nos leva ao erro de associar os epifenômenos e a atividade cerebral em uma relação de causalidade – o que de acordo com o epifenomenalismo é uma

---

<sup>12</sup> Termo utilizado por Descartes para se referir a mente (cf. Tripicchio, 2004, p. 25).

ilusão. A rigor, o epifenomenalismo é “um meio termo a que se chega entre o desejo de respeitar uma abordagem rigorosamente científica da explicação do comportamento e o desejo de respeitar o testemunho da introspecção” (Churchland, 2004, p. 32). Contudo, o apego ao testemunho da introspecção nos leva, necessariamente, a alguma versão de dualismo.

De acordo com Churchland (2004), o behaviorismo filosófico desponta como a primeira tentativa de resolução dos problemas do dualismo. Fortemente influenciado pelo positivismo lógico e usando uma metodologia calcada em uma análise cuidadosa da linguagem, os behavioristas procuram escamotear os problemas de ordem metafísica, rotulando-os como pseudoproblemas. Churchland (2004, p. 49) salienta que o behaviorista:

Afirma que falar sobre emoções, sensações, crenças e desejos não é falar sobre episódios espirituais interiores, mas um modo de falar sobre padrões de comportamento, potenciais ou reais. Em sua forma mais radical e simples, o behaviorismo filosófico afirma que toda sentença sobre estados mentais pode ser parafraseada, sem perda de significado, numa longa e complexa sentença sobre que comportamento observável iria resultar se a pessoa em questão estivesse nesta, naquela ou em outra circunstância observável.

Embora o behaviorismo, a princípio, tenha sido um alento em relação aos problemas do dualismo, ele também apresenta limitações. Seus dois principais problemas são: a incomoda negação categórica de nossos estados mentais – ter uma dor não parece se resumir ao estímulo, há uma experiência que se revela na introspecção que a identifica como desagradável; e a disposição<sup>13</sup> comportamental com múltiplas vias, uma lista de condicionais de caráter infinito ou excessivamente longa (cf. Churchland, 2004, p. 50-51).

A teoria da identidade<sup>14</sup> (o materialismo reducionista) é outra alternativa ao dualismo. Podemos descrevê-la de modo simples como a ideia de que “cada tipo de estado ou processo mental é numericamente idêntico a (é a mesma coisa que) algum tipo de estado ou processo mental no interior do cérebro, ou no sistema nervoso central” (Churchland, 2004, p. 53).

Não precisamos ir muito longe para perceber o problema do argumento da teoria da identidade. Basta a aplicação da lei de Leibniz e termos exposto o problema: para que dois itens sejam numericamente idênticos é necessário que todas as propriedades de um deles seja percebida também no outro – em termos lógicos teríamos  $(x) (y) [(x=y) \rightarrow (f) (fx \leftrightarrow fy)]$ . Como

---

<sup>13</sup> De acordo com o behaviorismo, a noção de disposição está associada as ações que um indivíduo realiza quando apresenta supostamente um estado mental do tipo “quero comer camarão”. Afirmar que Jessica quer comer camarão implica que: 1) Jessica diria que sim, se fosse indagada se gostaria de comer camarão; 2) ao lhe oferecerem lagosta e camarão, ela iria considerar primeiro a possibilidade de comer camarão; 3) ao ser informada que o camarão demoraria um pouco mais para ser servido, ela esperaria, e assim por diante.

<sup>14</sup> Como veremos, o funcionalismo dennettiano apresenta resquícios dessa corrente, pois apesar de vislumbrar outros “arranjos” físicos (orgânicos ou inorgânicos) o funcionalismo apresenta uma teoria da identidade mais branda ao conceber a correspondência entre *Espécime Mental* e *Espécime Físico*.

sugestão de argumento contra a teoria da identidade, essa lei fundamenta a tese segundo a qual se encontrarmos uma propriedade de um estado cerebral a qual não corresponde nenhum estado mental (ou vice-versa) a tese da identidade terá sido refutada (cf. Churchland, 2004, p. 59). Além disso, invariavelmente, a teoria da identidade, como quase todo argumento materialista, utiliza termos extraídos da psicologia popular<sup>15</sup> para considerar a mente como produto do cérebro – e, por assim dizer, oscila entre materialismo e dualismo.

Embora essa seja uma grande limitação de teorias calcadas no fisicalismo – conceber a mente ou estados mentais da perspectiva dualista, ou seja, baseado na introspecção – ainda assim essa corrente é considerada por muitos como a melhor maneira de se iniciar as pesquisas sobre a inteligência consciente, pois, diferentemente da abordagem dualista, procura se fundamentar nos trabalhos da neurociência, da psicologia cognitiva, da etologia e da ciência da computação. Essa aproximação das ciências empíricas proporciona uma vantagem aos teóricos do materialismo na edificação de teorias. Quanto à abordagem dualista, acusam-na de assumir uma postura negativista, ou seja, procuram descrever simplesmente o que a mente não é – se limitam a explicar a condição imaterial da mente. Sobre esse aspecto, Churchland (2004, p. 44) tece um comentário contundente:

O dualista pode nos dizer algo sobre a constituição interna da coisa mente? Sobre os elementos não-materiais que a constituem? Sobre as leis que governam seu comportamento? Sobre as conexões estruturais da mente com o corpo? Sobre o modo como ele opera? Ele pode explicar as capacidades e patologias humanas em termos de sua estrutura e defeitos? O fato é que o dualista não pode fazer nada disso, porque jamais foi formulada uma teoria detalhada sobre a coisa-mente.

O problema epistemológico é mais um ponto que gostaríamos de tratar aqui brevemente. Ele está associado ao modo pelo qual obtemos conhecimento, como suprimimos o ceticismo em relação às outras mentes, e pode ser dividido analiticamente em duas perspectivas: o problema das outras mentes e o problema da autoconsciência.

A novidade que surge com o problema das outras mentes pode ser descrita como a falta de justificativa das inferências que realizamos a partir da observação do comportamento das outras pessoas. Salvo a nossa própria experiência introspectiva, o que nos leva a crer que alguma outra entidade (um robô altamente sofisticado, um alienígena, uma inteligência artificial ou outro indivíduo humano) é um ser dotado de consciência? Uma primeira resposta seria o estabelecimento de uma correspondência entre generalizações psicológico-comportamentais e as generalizações empíricas padrão, ou seja, “se o comportamento do tipo ‘B’ é apresentado por uma criatura, então em geral um estado mental do tipo ‘S’ está ocorrendo” (Churchland,

---

<sup>15</sup> A *folk psychology*.

2004, p. 116). Embora implausível, a conclusão a que se chega é que, devido ao fato de podermos observar apenas os aspectos comportamentais da díade mente-comportamento, não se justifica a crença nas outras mentes. De tal forma que:

A crença em outras mentes exige inferências a partir do comportamento, essas inferências exigem generalizações sobre as criaturas em geral; essas generalizações, por sua vez, somente podem ser justificadas pela experiência das criaturas em geral; mas tudo o que podemos ter é a experiência do nosso próprio caso individual (Churchland, 2004, p. 117).

De acordo com Churchland (2004), há três soluções clássicas para o problema das outras mentes: 1) O argumento por analogia defende a ideia de que calcamos nossas inferências sobre as conexões psicológicas-comportamentais a partir da observação do nosso caso particular. Como a minha experiência demonstra que determinado estado mental está associado a um comportamento correspondente ou outro, ao observar em outro indivíduo um comportamento semelhante ao meu, concluo que o outro deve estar vivenciando o estado mental correspondente. Contudo, uma inferência baseada em um único caso particular de observação (o meu) não pode ser suficiente para estabelecer uma generalização. Trata-se de uma tese fundamentada no aspecto mais pálido do argumento indutivo – similar à afirmação de que todo animal marinho não possui pulmões apenas examinando anatomia de um único peixe. Talvez seja prudente questionar se “a nossa firme confiança na existência de outras mentes pode realmente ser justificada e explicada por um argumento assim fraco” (Churchland, 2004, p. 118). Além disso, identificamos os nossos próprios estados mentais a partir de uma rede conceitual preestabelecida pela psicologia popular. Portanto, os termos utilizados para realizar a analogia seriam delineados por outras mentes, comprometendo ainda mais as inferências indutivas (cf. Churchland, 2004, p. 119); 2) O problema das outras mentes também é abordado pelo behaviorismo. Trata-se de um modo para tentar escamotear o problema das outras mentes. Segundo essa abordagem, as generalizações que permeiam a relação psicológico-comportamental não deveriam receber o status de generalizações. Aquelas generalizações seriam interpretações das operações reduzidas a termos psicológicos neles contidos, de modo que dispensam a justificação empírica. Nesse sentido, a entidade que apresentar determinado comportamento, ou predisposição para tal, deve ser considerada um ente consciente. (cf. Churchland, 2004, p. 120); 3) Por fim, temos a hipótese de explicação. Essa abordagem pode ser descrita da seguinte maneira: se uma teoria apresenta uma hipótese que pode ser associada a circunstâncias observáveis e que posteriormente se revelam verdadeiras, estamos diante de uma justificação hipotético-dedutiva. Uma teoria (da mente) deve ser creditada, uma vez que

ela explica e realiza previsões acerca de fenômenos não observáveis, com o auxílio de sua fundamentação em fenômenos observáveis previstos (cf. Churchland, 2004, p. 122).

Uma visão contemporânea muito debatida é a autoconsciência<sup>16</sup>. Ela caracteriza-se por graus de refinamento da percepção através do aprendizado – se nos especializamos em determinada área adquirimos uma maior consciência, um maior domínio sobre um assunto específico, uma visão mais ampla sobre o recorte do real (cf. Churchland, 2004, p. 125). Podemos concebê-la como uma concepção evolucionária, pois:

É evidente que a percepção, quer interna quer externa, é substancialmente uma habilidade apreendida. A maior parte desse aprendizado ocorre em nossa tenra infância, é evidente: o que é agora perceptualmente óbvio para nós era uma discriminação sutil aos oito meses. Mas sempre há lugar para se aprender mais (Churchland, 2004, p. 126).

A autoconsciência se mostra, de acordo com a concepção evolucionista, dependente da evolução do cérebro. Nesse sentido, ela é secundária em relação à percepção do ambiente (cf. Churchland, 2004, p. 129).

Apesar de apresentamos alguns aspectos que permeiam a investigação do problema da natureza da consciência, apenas resvalamos nas dificuldades que emergem à abordagem do tema. O pensamento dennettiano nos ajudará a nos aprofundarmos um pouco mais, pois, buscaremos dirimir os conflitos que envolvem uma teoria contra-intuitiva, já que “alguns dos paradoxos mais desconcertantes da consciência surgem porque nos apegamos por muito tempo a um bom hábito de pensar, um hábito que geralmente nos mantém longe de problemas” (Dennett, 1991a, p. 101 – nossa tradução).

Na abordagem dennettiana, observamos uma dura crítica ao dualismo cartesiano que se estende para outras abordagens de mesma natureza. No entanto, não é lícito rotulá-lo de materialista reducionista – o que, em nosso entendimento, coaduna-se com a interpretação de que o seu pensamento se caracteriza por ser uma terceira possibilidade de abordar o problema da consciência (cf. Teixeira, 2008, p. 08), reportando a uma filosofia de caráter funcionalista<sup>17</sup> cujo caráter especial assume a forma de um funcionalismo homuncular – um pensamento que admite que os dados comportamentais podem ser tomados como estados psicológicos que são explicados a partir da concepção do sujeito como sistemas de componentes interligados, haja vista que “cada componente é um homúnculo, no sentido de que é pensado como um pequeno agente ou burocrata operando dentro do sujeito que o contém” (Bunnin; Tsui-James, p. 2002,

---

<sup>16</sup> Um contraponto à concepção tradicional – de caráter empirista ortodoxa – que defende a tese que as percepções interna e externa são indistinguíveis. A primeira pode se mostrar falsa, enquanto que a segunda é autoevidente (cf. Churchland, 2004, p. 127).

<sup>17</sup> Também é rotulada pelo autor de teoria empírica da consciência.

176) –, e que, ao mesmo tempo que tece uma crítica, estabelece uma ampliação do funcionalismo mecânico de Jerry Fodor.

De acordo com Dennett (2006), há uma distinção, por exemplo, entre aquilo que o olho do leão diz para seu cérebro e aquilo que o olho do leão diz para si em uma representação. No entanto, segundo o filósofo, Fodor parece desconsiderar esse fato. Uma psicologia bem-sucedida, a neocognitivista<sup>18</sup>, seria aquela que admite um sistema interno de representação. Essa é a interpretação que o filósofo faz de Fodor.

De certo, não é possível a representação do nada, ou seja, toda representação é sempre uma representação de algo para alguém, bem como toda representação (ou sistema de representação) necessita de pelo menos um administrador do sistema. Nesse sentido, o neocognitivismo de Fodor também exige a existência de um agente, ou vários burocratas, internos e apartados do sistema, uma espécie de homúnculo não-eliminado. Sabemos que toda tese psicológica fundamentada em homúnculos não-elimináveis está fadada ao erro, pois recai em um processo de regressão ao infinito (cf. Dennett, 2006, p. 154). Embora seja uma concepção que esteja vinculada à processos computacionais – o que leva a tese de Fodor a uma aproximação ao pensamento dennettiano –, essa postura radical de representação estiola a tese neocognitivista. Dennett o acusa de assumir um compromisso com a hipótese infundada segundo a qual somente uma representação explícita é uma representação (cf. Dennett, 2006, p. 157). Por ora, vejamos o que Dennett (1991a, p. 14 – nossa tradução) diz a respeito de sua concepção de homúnculo:

[...] as teorias que postulam esses homúnculos ("homenzinhos" no cérebro) nem sempre devem ser evitadas, mas sempre que homúnculos são chamados para ajudar, é melhor que sejam funcionários relativamente estúpidos – não como os brilhantes dramaturgos freudianos que deveriam produzir novas cenas de sonho todas as noites para cada um de nós! O modelo que estamos considerando elimina totalmente o dramaturgo e conta com o "público" (análogo àquele que é "isso" no jogo de festa) para fornecer o conteúdo. O público não é um idiota, é claro, mas pelo menos não precisa ter uma teoria de suas próprias ansiedades; só precisa ser orientado por eles para fazer perguntas.

O funcionalismo dennettiano veio em resposta ao problema da introspecção e ao complexo problema das outras mentes – é uma teoria de caráter não-reducionista. Em virtude

---

<sup>18</sup> Apesar de Fodor, segundo Dennett (2006), delinear um tipo de teorização que denomina de psicologia cognitiva, percebemos que ela é melhor descrita sob o termo de neocognitivismo, por apresentar como a marca "distintiva de sua teorização [...] a utilização sem receio de uma caracterização intencional de processos e de análise *intelectualistas* da percepção e de outros processos cognitivos em termos de fluxo de informação, testes de hipóteses, inferências e tomada de decisão" (Dennett, 2006, p. 143 – grifo nosso). Uma abordagem nesses parâmetros, de acordo com o filósofo, reúne a produção recente dos trabalhos da psicologia, linguística, inteligência artificial e de áreas específicas da filosofia contemporânea.

disso, leva em conta os trabalhos das ciências empíricas, sem abrir mão da contribuição da filosofia.

Considere o seguinte caso: se perguntarmos para um físico qualquer “como podemos conceber a existência de algo?” a resposta possivelmente seria “a partir da possibilidade de sua mensuração”. Essa resposta é interessante e problemática. Interessante porque é a resposta de uma das áreas de conhecimento mais respeitadas do nosso tempo. O problema, no entanto, reside no momento em que realizamos uma introspecção e constatamos a existência de nossos próprios estados mentais. Não podemos, simplesmente, admitir a negação dessas emoções, nem tampouco a sua mensuração é possível – seria absurda, para psicologia popular, a afirmação, por exemplo, de que agora estou com minha alegria no nível 8,5 e se eu baixar um pouco mais minha tristeza talvez possa elevá-la para 9,0 –, pois as qualidades não são passíveis a mensuração. Ao que parece, as contribuições das ciências empíricas também têm suas limitações, porquanto não podemos reduzir estados mentais às quantidades, bem como submeter hipóteses nesse campo a testes experimentais.

Observando essas limitações, os filósofos (os teóricos da mente inclusive) aplicam um método não empírico – seja para refutar ou corroborar uma hipótese –, a saber: os experimentos de pensamento (EPs). Esse método também será objeto de investigação neste capítulo. Assim, nos dedicaremos ao exame sobre a aplicação e importância desse tipo de experimento e de como Dennett aborda essa questão.

### **1.1 As estranhas inversões: Darwin e Turing**

Um dos pressupostos fundamentais do pensamento dennettiano é o evolucionismo – por explicar com elegância e simplicidade a complexidade dos projetos da natureza. Conquanto outros pesquisadores<sup>19</sup> tenham teorizado sobre a evolução, a supremacia de Darwin e Wallace está em descrever, de modo plausível, um mecanismo pelo qual a evolução ocorre, a saber: a seleção natural. A simplicidade da ideia é tal que podemos estendê-la para outros domínios, inclusive à psicologia. Os contornos da teoria da evolução por seleção natural são claros. Dennett (1991a, p. 200 – nossa tradução) nos traz suas características:

- (1) Variação: uma abundância contínua de diferentes elementos;
- (2) Hereditariedade ou replicação: os elementos têm a capacidade de criar cópias ou réplicas de si mesmos;
- (3) “Adequação” diferencial: o número de cópias de um elemento que são criadas em um determinado tempo varia, dependendo das interações entre as características desse elemento (seja o que for que o torne diferente de outros elementos) e as características do ambiente em que persiste.

---

<sup>19</sup> Os trabalhos de Jean de Lamarck (1744-1829) são um bom exemplo.

A teoria darwiniana, segundo Dennett (2009a), é uma das mais revolucionárias e assustadoras do pensamento humano. A novidade que surge com Darwin é a tese de que é possível existir adaptação sem que para isso haja um artífice consciente, ou seja, a ignorância absoluta seria o “arquiteto”, de modo que uma das consequências fundamentais da corrida evolutiva é que “para criar uma máquina perfeita não é necessário saber como criá-la” (Dennett, 2009a, p. 61 – nossa tradução). Essa proposição, claramente contra-intuitiva, irá, em uma análise atenta, apresentar resumidamente o propósito especial da teoria e expressar em poucas palavras todo o significado de Darwin, que através de uma inversão estranha de raciocínio, parece considerar a ignorância absoluta totalmente capaz de substituir a ideia de uma sabedoria absoluta, em todas as realizações da capacidade criativa. Esse argumento servirá como um dos pilares para a elaboração de uma teoria da consciência efetivamente calcada nas ciências empíricas, ensejando a teoria memética<sup>20</sup>.

O pensamento anterior ao darwinismo, rotulado de teoria da criação lenta, era calcado em uma visão do universo segundo o qual o mental vem primeiro. Essa tese imperava sem contestação. Em termos concretos, isso significa que se acreditava firmemente que um Deus inteligente seria a origem de todo *design*, a respostas para todos os “porquês” (cf. Dennett, 2009a, p. 61). O filósofo John Locke (1632-1704) compartilha desse entendimento, ao descrever uma consciência divina (ou originária) – construindo um argumento que nega a possibilidade da geração de inteligência a partir da organização material ignorante. Isso significa que a mente surgiria primeiro – ou no máximo concomitantemente aos demais entes. Afastando-se da hipótese de que a mente poderia emergir em alguma data subsequente, o filósofo inglês não acredita que seja algum efeito de convergências fenomenais mais simples e de nível irracional. Essa postura se apresenta como a justificativa, de caráter lógico-matemático, de uma cosmogonia judaico-cristã: no princípio havia um ser que possuía uma mente (o ser cognitivo lockeano). O novo que surge com Locke é que aquela entidade originária pensante e racional da tradição (o arquiteto e engenheiro do universo) seria agora reapresentado por um novo selo, a autenticação científica – uma vez que estando atrelado a um teorema matemático, estaria supostamente imune a qualquer refutação (cf. Dennett, 1995a, p. 27-28).

Por outro lado, David Hume (1711-1776), ao expor os problemas com essa tese, chegando mesmo a se aproximar de solução oferecida por Darwin, admitiu que não seria possível tratá-la com seriedade (cf. Dennett, 1995a, p. 32-33). Ele desenvolveu bons argumentos sobre como é possível as impressões e ideias fundirem-se ontologicamente em um

---

<sup>20</sup> A teoria memética, bem como o modelo *pandemonium* dennettiano, será tratado no próximo capítulo.

tipo de desdobramento análogo ao da ligação química, permitindo o surgimento dos hábitos na mente humana. No entanto, esses argumentos não poderiam ser submetidos aos testes das ciências empíricas (cf. Dennett, 1997, p. 69).

Com efeito, é com a pugnacidade de Darwin que efetivamente iremos poder lançar mão a uma resposta satisfatória para o problema. Sua judiciosa tentativa de demonstrar que o processo de evolução prescinde a qualquer mentalidade originária, rompe as fronteiras da biologia. Dennett (1995a, p. 39) realiza uma observação da sua importância na seguinte passagem:

a mutação darwiniana, que a princípio parecia ser apenas uma nova maneira de pensar sobre os tipos de biologia, pode se espalhar para outros fenômenos e outras disciplinas, [...] Existem problemas persistentes dentro e fora da biologia que se dissolvem prontamente quando adotamos a perspectiva darwiniana sobre o que faz com que um tipo de coisa seja o que é, mas a tradição da tradição para essa ideia persiste.

A proeminência teórica do darwinismo, em detrimento ao pensamento tradicional, reside nas acuradas características identificadas pelo biólogo no desdobramento das espécies – embora não utilizasse um termo específico para o fenômeno em questão. Havia uma dificuldade de estabelecer uma descrição explícita na ausência de tal termo. Podemos englobar esses recursos ao canalizá-los em um único termo, a saber: o algoritmo<sup>21</sup>. Esse conceito não é uma descoberta darwiniana, sua aplicação é percebida, já nos tempos que precederam à Darwin, em outras áreas de atividade – como, por exemplo, na resolução de uma divisão extensa ou para estabelecer a ordem alfabética em uma lista de palavras. Nesse sentido, podemos defini-lo como “um certo tipo de processo formal com o qual se pode contar – logicamente – para produzir um certo tipo de resultado sempre que for *executado* ou instanciado” (Dennett, 1995a, p. 50 – grifo nosso). O novo que surge com a concepção darwiniana é sua notória influência nas reflexões de matemáticos e lógicos no que tange à ontologia dos algoritmos e a possibilidade de sua aplicação. Esses estudos culminaram, nas primeiras décadas do século XX, no surgimento do computador digital – o que implicou na necessidade de uma compreensão ainda mais aprimorada dos algoritmos, numa perspectiva ampliada de áreas de aplicação.

Não obstante, consoante Dennett (1995a), Darwin não delineou categoricamente um algoritmo específico (um subgênero), desvelou na verdade uma grande classe de algoritmos interrelacionados. Todavia, é possível reformular a sua tese fundamental: a vida floresceu em nosso planeta, há bilhões de anos, a partir de uma única raiz (da chamada árvore da vida) por um processo algorítmico de “um ou outro”.

---

<sup>21</sup> O termo tem origem no nome de Mohammad al-Khuwarizmi, um matemático e astrólogo árabe do século IX que contribuiu de modo determinante para introduzir o sistema de numeração indiano no ocidente.

Conforme Dennett (1995a), ao refletir sobre os algoritmos, a maioria dos teóricos canalizam sua atenção para um modelo que não abrange as propriedades cujo filósofo vislumbra – quando um matemático abraça esse conceito, intenciona aplicar um modelo que pode ser submetido à comprovação e uma posterior aplicação no cálculo de funções matemáticas de seu interesse<sup>22</sup>. Mas o tipo de algoritmo que o interessa se distancia dos sistemas numéricos, bem como de outros objetos da matemática. Em termos concretos, isso significa que, para o pensamento dennettiano, o modelo de algoritmo que interessa é aquele que possibilita classificar, vencer e construir coisas. Dado que grande parte dos debates matemáticos sobre algoritmos está vinculado ao alcance dos seus poderes e se são matematicamente comprovados, as pessoas estão inclinadas ao erro de entender que um processo algorítmico está isento da atividade do acaso ou aleatoriedade – na verdade, numa divisão longa, por exemplo, está inserido um quinhão de aleatoriedade (cf. Dennett, 1995a, p. 52). O autor ainda traça uma importante linha tangencial entre a divisão longa e aleatoriedade:

O divisor entra no dividendo seis ou sete ou oito vezes? Quem sabe? Quem se importa? Você não precisa saber; você não precisa ter inteligência ou discernimento para fazer uma longa divisão. O algoritmo direciona você apenas para escolher um dígito – aleatório, se quiser – e conferir o resultado. Se o número de matriz escolhido for muito pequeno, aumente-o em um e comece de novo; se muito grande, diminua. O bom da divisão longa é que sempre funciona, mesmo que você seja muito estúpido ao fazer sua primeira escolha; nesse caso, leva um pouco mais de tempo (Dennett, 1995a, p. 53).

Nessa passagem, Dennett reverbera que a aleatoriedade é um aspecto inerente ao processo algorítmico. O acaso – a irracionalidade – também é responsável pelo comportamento inteligente, ou seja, a inteligência tem certo grau de ignorância. Ficamos maravilhados ao observar que computadores são capazes de realizar tarefas complexas a partir da absoluta ignorância. Nos perguntamos: como é possível uma máquina ignorante fazer algo tão inteligente? Se regularmos nosso pensamento para aplicação de um processo aleatório de seleção de candidatos, que são testados de modo mecânico, chegamos aos algoritmos interessantes. Portanto, não só há uma interferência da presença do acaso nos poderes dos algoritmos, mas, segundo o filósofo, em alguns casos eles são condições de uso desses poderes (cf. Dennett, 1995a, p. 53). Dennett (*Ibidem*) realiza uma observação da sua importância na seguinte passagem:

Podemos começar a nos concentrar no filo dos algoritmos evolutivos, considerando algoritmos cotidianos que compartilham propriedades importantes com eles. Darwin chama nossa atenção para ondas repetidas de concorrência e seleção; portanto, considere o algoritmo padrão para organizar um torneio de eliminação, como um

---

<sup>22</sup> O exemplo que Dennett nos dá é a “divisão longa [...] um procedimento para quebrar um grande número em seus fatores primos é aquele que atrai atenção no mundo exótico da criptografia” (Dennett, 1995a, p. 52).

torneio de tênis, que eventualmente culmina com quartas de final, semifinais e depois uma final, determinando o vencedor solitário.

Os fãs de futebol – mais precisamente aqueles que são mais familiarizados em copa do mundo – podem assimilar com maior idoneidade o exemplo a seguir: é comum que nesse tipo de evento esportivo (a exemplo de vários outros esportes) que as fases de eliminação obedeçam a um chaveamento seletivo de caráter algorítmico. As últimas edições da copa, por exemplo, apresentaram um total de 32 seleções nacionais, divididas em oito grupos de quatro equipes que jogam entre si – passando para a segunda fase apenas as duas melhores equipes de cada grupo. Essa é a configuração da primeira fase. Na figura 1.1 temos a simulação dos selecionados para a segunda fase:



Figura 1.1

Segunda fase da copa do mundo de 2022 (Fonte: Fifa.com)

Ao estabelecer uma simulação, em um computador que esteja ligado à internet, as atualizações podem ser realizadas em tempo real, ou seja, na medida em que os jogos forem finalizando. Você poderia realizar essa tarefa em um quadro com giz. As equipes perdedoras ficam pelo caminho e qualquer um seria capaz de conduzir as equipes vencedoras para a fase seguinte ao final de cada partida (seja identificando ou apagando o nome dos derrotados).

De acordo com Dennett (1995a), há um amálgama de habilidade e acaso, uma fusão natural e inevitável em toda concorrência real à medida que possam ter uma ampla variação. Uma partida de copa do mundo em um estádio com o gramado comprometido (buracos, murrinhos-artilheiro, etc.) eleva o grau de contingência, bem como uma disputa por pênaltis após uma partida desgastante acrescida de prorrogação. Todavia, ainda que seja inegável a

presença do acaso nesse processo, há uma forte tendência, observada estatisticamente, das equipes melhores preparadas avançarem neste tipo de competição. Assim, a capacidade de uma competição de identificar, a longo prazo, as discrepâncias entre as habilidades dos participantes é dirimida (e não reduzida à zero) pelo fator aleatoriedade. Este preâmbulo de esclarecimento conceitual é factual e também “verdadeiro para os algoritmos evolutivos na natureza dos torneios de eliminação no esporte, às vezes é negligenciado pelos comentaristas sobre a evolução” (Dennett, 1995a, p. 55 – nossa tradução). Sobre a perigosa ideia de Darwin, Dennett (1995a, p. 60 – nossa tradução) nos traz um panorama que facilita o entendimento de seu pensamento:

o nível algorítmico é o nível que melhor explica a velocidade do antílope, a asa da águia, a forma da orquídea, a diversidade de espécies e todas as outras ocasiões por admiração no mundo da natureza. É difícil acreditar que algo tão irracional e mecânico como um algoritmo pode produzir coisas tão maravilhosas. Não importa quão impressionantes os produtos de um algoritmo, o processo subjacente sempre consiste em nada além de um conjunto de etapas irracionais individualmente, que se sucedem sem a ajuda de qualquer supervisão inteligente; eles são "automáticos" por definição: o funcionamento de um autômato.

Esse tipo de processo exige uma retroalimentação entre os envolvidos – com chances cegas, ou seja, não há um burocrata que realiza escolhas – e nada além disso. Os algoritmos com os quais temos maior familiaridade apresentam produtos mais simples, ou seja, elaboram “listas de divisão longa ou alfabetizadas ou descobrem a renda do contribuinte médio” (*Ibidem*). Os que apresentam maior complexidade e sofisticação são responsáveis pelos fantásticos produtos que nos são apresentados cotidianamente (aqueles que simulam mundos virtuais completos, povoados de seres nunca vistos antes). No entanto, o mundo natural é ainda muito mais complexo, expresso por vários níveis de magnitudes. Seria possível que a biosfera fosse resultado de uma enxurrada de processos algorítmicos que se nutrem de acaso? E se sim, quem seria o projetista? Absolutamente ninguém. Segundo a teoria darwiniana<sup>23</sup>, seria o resultado de um processo cego e algorítmico.

Diante disso, seguindo a esteira de Dennett (2009a), na teoria darwiniana, a ignorância absoluta é o grande arquiteto. Sem excrescências, podemos afirmar categoricamente, respaldados por esse princípio fundamental de todo sistema, que é simplesmente possível realizar a construção de uma máquina eficaz, sem com isso precisar estabelecer uma inteligência apriorística. Assim, podemos resumir a contribuição da teoria evolutiva à teoria da consciência dennettiana: há o oferecimento de uma estranha inversão de raciocínio que considera a ignorância absoluta como “totalmente qualificada para substituir a sabedoria

---

<sup>23</sup> Segundo Dennett (1995a), o próprio Darwin declarou (em uma carta ao geólogo Charles Lyell) que se sua teoria da evolução necessitasse de um artífice inteligente – tingida de argumentos miraculosos – seria digna de rejeição.

absoluta em todas as realizações da habilidade criativa" (Dennett, 1995a, p. 65 – nossa tradução).

Essa estranha inversão, de acordo com Dennett (1995a), é uma excelente maneira de refutar, de modo definitivo, a tese de Locke. Contudo, a ideia de que a mente possa ser o efeito e não a causa da evolução é (para muitos) uma ideia altamente subversiva – há uma dificuldade natural de acomodar esse entendimento em suas mentes. Talvez seja possível dirimir esse preconceito expondo algumas questões: o que condicionaria um subproduto, oriundo de forças irracionais e sem propósito, a receber o rótulo de vulgar? Por que a entidade mais relevante de todas não poderia emergir de algo irrelevante? Por que tudo que apresenta uma excelência deve ser uma dádiva divina? A inversão realizada pelo darwinismo alvitra que deixemos de lado essa postura presunçosa e vislumbremos um tipo de “entidade” cujo valor e propósito podem perfeitamente surgir de forças ignorantes (cf. Dennett, 1995a, p. 65-66).

No século XX, a inversão estranha, apontada por Dennett (2009a) na teoria darwiniana, é percebida na teoria computacional de Alan Turing. Nos velhos tempos os “computadores” precisavam saber matemática – eram matemáticos bem treinados –, mas Turing percebeu que isso não era necessário, ou seja, os computadores não precisam reconhecer as razões. Para entender a concepção de mente dennettiana é preciso abordarmos o que os pesquisadores de IA, em sua época, concebiam como sendo a natureza do mental, pois o filósofo retoma essa ideia (cf. Teixeira, 2008, p. 08). O conceito de pensamento (concebido como comportamento linguístico) de Turing é pragmático e operacional – o mesmo entendimento que encontramos no pensamento dennettiano. Para ser uma máquina de computador perfeita não é necessário saber o que é aritmética (cf. Dennett, 2009a, p. 61).

A concepção de que o algoritmo permite o desenvolvimento de processos mecânicos é anterior à Turing. Entrementes, os trabalhos realizados pelos matemáticos (dos quais podemos destacar Kurt Gödel, Alonzo Church e o próprio Alan Turing), nas primeiras décadas do século passado, nos legaram à compreensão atual do termo. Segundo essa compreensão, há três características principais dos algoritmos, relativamente complicados de assimilar<sup>24</sup>, que são relevantes para suas pretensões. Primeiro, temos a *neutralidade do substrato* – na realização de uma divisão longa a eficácia do resultado independe do material utilizado (lápiz ou caneta, papel ou pergaminho), para efetivar um processo algorítmico basta um sistema qualquer de símbolos que permitam a execução das etapas pré-estabelecidas – é a estrutura lógica que garante o poder do procedimento. Na sequência, temos a *falta de sentido subjacente* – apesar do *design*

---

<sup>24</sup> Consoante Dennett (1995a), ainda hoje essas características são pivôs de controvérsias.

proporcionar resultados excelentes e de aparente complexidade, cada fase do procedimento, bem como a transição entre elas, são de caráter simplório, ou seja, são formados por etapas mais simples do processo – simplório o suficiente para que seja exequível por um estúpido ou dispositivo mecânico direto. Por fim, temos que *garantir o resultado* – não importa o que está sendo executado por um algoritmo, ele o realizará em uma economia que garante o resultado esperado (cf. Dennett, 1995a, p. 51).

Nessa perspectiva, um programa de computador é formado por um tipo ou outro de algoritmo, um sistema composto por uma série de fases com níveis de complexidade inferior ao todo. A supremacia dos circuitos eletrônicos é um fato (dada a sua velocidade de processamento), mas isso não quer dizer que seu desenvolvimento se resume a um tipo de estrutura canalizada nas peculiaridades causais dos elétrons que disparam nos chips de silício. Assim, é de se esperar que “os mesmos algoritmos podem ser realizados (ainda mais rápido) por dispositivos que desviam fótons em fibras de vidro, ou (muito, muito mais lento) por equipes de pessoas que usam papel e lápis” (Dennett, 1995a, p. 51 – nossa tradução).

Mais adiante, veremos como o funcionalismo homuncular dennettiano explora essa noção de algoritmo. Decerto, as capacidades dessas máquinas de realizar procedimentos algorítmicos, com extrema velocidade e alto grau de confiança, permite ao filósofo defender com mais afinco os aspectos contra-intuitivos de sua teoria por meio de uma analogia entre cérebros e computadores.

## **1.2 O papel da ciência cognitiva e da I.A.**

De acordo com Gardner (1996), em meados do século IV a.C., com a edificação da academia de Platão (428/27-348/47a.C.) – que visava difundir o pensamento socrático –, são iniciadas as investigações acerca do conhecimento. A perspectiva platônica de conhecimento estava intimamente ligada à matemática e às ciências exatas. Por conseguinte, o conhecimento por excelência seria sobre as formas idealizadas, alcançado exclusivamente pelo intelecto, de tal modo que, através dos sentidos, só poderíamos vislumbrar suas cópias pálidas. O interesse de compreender a apreensão do conhecimento genuíno já nos dominava desde os primeiros trabalhos filosóficos da civilização grega. Essa concepção está bem retratada no diálogo Menon, na clássica passagem em que Sócrates (469-399 a.C.) instrui um escravo na resolução de um problema de geometria – segundo Platão não haveria ensinamento, mas uma canalização das ideias até a consciência (cf. Gardner, 1996, p. 17).

A inclinação pela busca do entendimento da natureza do conhecimento se expandiu pela tradição intelectual do ocidente: no medievo, a retomada do pensamento aristotélico, por

exemplo, que apresenta uma nítida aversão pelas ideias inatas, tem forte influência na escolástica; as ciências empíricas ganham destaque no período que vai do renascimento ao iluminismo – a teoria do conhecimento fica por conta de nomes como Descartes, Locke, Hume e Kant (1724-1804); mas é apenas nos anos finais do século XIX que se acentua a emergência de novas áreas de pesquisas – a psicologia, a linguística, a antropologia, a sociologia e as neurociências que assumem papel importante na investigação sobre a natureza da mente humana (cf. Gardner, 1996, p. 18).

Atualmente, munidos com um aparato tecnológico de ponta, e auxiliados por novos conceitos, a ciência cognitiva se apresenta como uma das mais importantes áreas de investigação para as pesquisas relacionadas à natureza da inteligência consciente. Sobre esse ponto Gardner (1996, p.18-19) assevera – e nos explica melhor:

Assim como seus antigos colegas, os cientistas cognitivos de hoje perguntam o que significa conhecer algo e ter crenças precisas, ou ser ignorante ou estar errado. Eles procuram entender o que é conhecido – os objetos e sujeitos do mundo externo e as pessoas que conhece – seu aparelho perceptivo, mecanismo de aprendizagem, memória e racionalidade. Eles investigam as fontes do conhecimento: de onde vem, como é armazenado e recuperado, como ele pode ser perdido? Eles estão curiosos com a diferença entre indivíduos: quem aprende cedo ou com dificuldade: o que pode ser conhecido pela criança, pelo cidadão de uma sociedade não letrada, por um indivíduo que sofreu lesão cerebral, ou por um cientista maduro?

E continua:

Além disso, os cientistas cognitivos, novamente como os gregos, conjecturam a respeito dos vários veículos do conhecimento: o que é uma forma, uma imagem, um conceito, uma palavra; e como estes 'modos de representação' se relacionam entre si? Eles se perguntam sobre as prioridades de órgãos sensoriais específicos em contraposição a uma 'compreensão geral' ou 'senso comum' central. Eles refletem acerca da linguagem, observando o poder e as armadilhas trazidos pelo uso das palavras e a sua possível influência predominante sobre pensamentos e crenças. E especulam extensivamente a respeito da natureza da própria atividade de conhecer: por que nós queremos conhecer, quais as restrições do conhecimento, e quais são os limites do conhecimento científico sobre o conhecimento humano? (*Ibidem*).

Entretanto, de acordo com Gardner (1996), a ciência cognitiva não está em consonância com os gregos em todos os aspectos. Os métodos empíricos, os experimentos, são os recursos utilizados para colocar à prova suas hipóteses. Seu desenvolvimento é alavancado pela emergência de novas disciplinas e de questões mais provocativas – como, por exemplo, se é possível ou não o desenvolvimento de máquinas pensantes (cf. Cescon, 2013, p. 161).

No final dos anos 40, com o intuito de debater a clássica questão acerca de como o sistema nervoso causa o comportamento, foi realizado um congresso<sup>25</sup> sobre os "mecanismos cerebrais do comportamento". Embora fosse apenas um congresso, dentre tantos outros

---

<sup>25</sup> Sob o título de Simpósio de Hixon, esse congresso, sediado no *California Institute of Technology*, reuniu representantes de várias áreas de pesquisas (cf. Gardner, 1996, p. 27-28).

realizados por cientistas de inclinação cognitivista, ele foi de grande relevância por dois motivos: primeiro, por estabelecer uma aproximação analítica entre o cérebro e o computador; segundo, por submeter o behaviorismo, e alguns dos seus pressupostos teóricos, a uma crítica contundente<sup>26</sup> (cf. Gardner, 1996, p. 43).

Segundo Gardner (1996), a ciência cognitiva foi desenvolvida com base nos seguintes pressupostos teóricos. A formalização, por Alan Turing, do conceito de *máquina* (a chamada máquina de Turing) foi extremamente relevante para as pesquisas sobre máquinas computadoradas. O *modelo neuronal* desenvolvido por Walter Pitts (1923-1969) e Warren MacCulloch (1898-1969) que demonstrou que as operações das células nervosas e suas ramificações poderiam ser traduzidas em termos lógicos. Essa descoberta proporcionou a aplicação da ideia de máquina de Turing à neurologia, possibilitando a concepção de um sistema nervoso formado por um sem números de neurônios *tudo-ou-nada* e uma máquina computadorada com competência para realizar operações diversas de caráter inequívoco. As *síndromes neuropsicológicas* – a segunda guerra mundial patrocinou uma série de casos de limitações cognitivas e patologias oriundas de danos cerebrais (nas descobertas intercambiadas de vários estudiosos, foi revelado a possibilidade de haver um padrão de regularidades na distribuição das habilidades cognitivas do sistema nervoso, em detrimento da concepção ambiental dos processos mentais), tais estudos delineiam os horizontes das pesquisas que tentam descrever como a mente se organiza. *A teoria da informação*: Claude Shannon (1916-2001) apresentou uma ideia à frente de seu tempo. Ele defendeu a tese de que circuitos elétricos – similares aos do computador – podem apresentar operações fundamentais de pensamento. Junto com Warren Weaver (1894-1978) trabalhou no desdobramento da tese central da teoria da informação – a informação pode ser criada sem a necessidade de nenhum conteúdo ou tema específico, meramente pela escolha entre duas opções que possuem o mesmo nível de plausibilidade.

Com efeito, a ciência cognitiva é "um esforço contemporâneo, com fundamentação empírica, para responder questões epistemológicas de longa data – principalmente aquelas relativas à natureza do conhecimento, seu desenvolvimento e seu emprego" (Gardner, 1996, p. 43) e se caracteriza por apresentar cinco aspectos basilares: 1) as *representações* – a ciência cognitiva é calcada na ideia de que é preciso estabelecer um nível de análise à parte, denominado de nível de representação, calcado em entidades representacionais (símbolos,

---

<sup>26</sup> Diante dos estudos desenvolvidos pelas ciências cognitivas, o behaviorismo se viu obrigado a rever seus conceitos, sob pena de ser completamente abandonado.

regras, imagens) e no modo pelo qual elas se relacionam; 2) o *computador* – é ao mesmo tempo um importante recurso para o desdobramento dos trabalhos da ciência cognitiva e modelo de reprodução<sup>27</sup> da mente humana; 3) a ausência *da emoção, do contexto, da cultura e da história* – esses elementos são preteridos pelos pesquisadores da ciência cognitiva com o intuito de torná-la viável; 4) *estudo interdisciplinar* – há uma inclinação para agregação dos trabalhos de outras áreas de pesquisas; 5) *raízes em problemas filosóficos clássicos* – é o fundamento lógico, o pressuposto das investigações da ciência cognitiva (cf. Gardner, 1996, p. 43). A figura 1.2 estabelece a arquitetura das relações interdisciplinares.

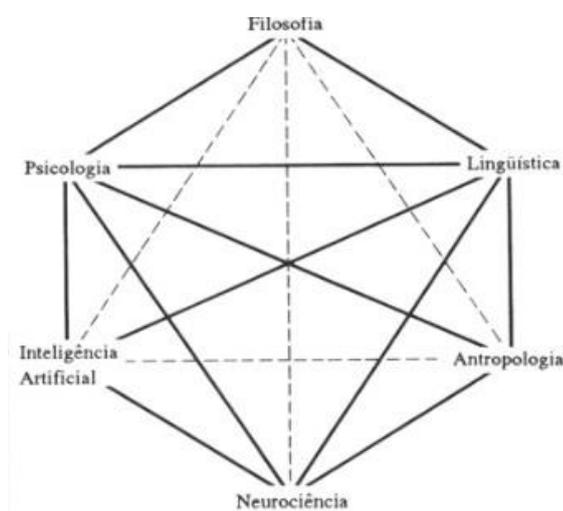


Figura 1.2

**Linhas Cheias** – fortes vínculos interdisciplinares.

**Linhas Tracejadas** – fracos vínculos interdisciplinares.

Fonte: Google

Dennett (1991a; 2012) acredita que munidos apenas de conceitos e métodos da filosofia da mente tradicional não é possível solucionar o problema da coisa-mente – ela continuaria sendo um mistério. Com eles chegaríamos, na melhor das hipóteses, a um ceticismo a níveis próximos aos de McGinn<sup>28</sup> e Nagel. De acordo com o filósofo, é trilhando o caminho das ciências cognitivas que podemos superar essa limitação. Sobre esse ponto, Dennett (1991a, p. 01 – nossa tradução) expressa sua opinião de forma clara e vigorosa:

O ritmo das descobertas está acelerando, graças à montanha de dados fornecidos pelas novas tecnologias de imagem cerebral, mas ainda mais às simulações de computador que expandiram e disciplinaram nossa imaginação, ampliando drasticamente o espaço

<sup>27</sup> A I.A. é a disciplina principal da ciência cognitiva (cf. Gardner, 1996, p. 43).

<sup>28</sup> Dennett (1991a; 1991b) critica veementemente a posição cética de Colin McGinn e Thomas Nagel sobre os limites do cérebro para lidar com a análise da consciência. Segundo o autor, McGinn esquece que esses limites estão sendo gradativamente superados pelos desenvolvimentos de novas tecnologias que permitem ampliar nossa percepção – além de proporcionar experimentos mais eficazes.

lógico dos modelos que podem ser investigados. Agora, podemos considerar seriamente hipóteses que, alguns anos atrás, eram simplesmente intransponíveis – "inconcebíveis", um filósofo poderia ter sido tentado a dizer. Esses poderes expandidos por computador estão sendo vigorosamente explorados por uma nova geração de teóricos e experimentalistas.

Atualmente, as pesquisas no campo da ciência cognitiva estão relacionadas à explicação de como neurônios – ou redes neurais – podem promover (seja em frequência de 40 Hz ou de tantas outras hipóteses) a inteligência consciente (cf. Cescon, 2013, p. 161).

Segundo Dennett (1991a; 1997; 2006), é possível que as mentes artificiais possam ser concretizadas a partir de qualquer coisa – ou seja, não se restringem ao material orgânico que constitui o cérebro humano. Ao obtermos um amplo domínio do entendimento da estrutura do mental (o que proporciona a percepção das dores, das crenças, etc.), deveríamos ter a competência de reproduzir mentes (e partes mentais) a partir de materiais diversos. Essa concepção – de que o papel das mentes é processar informação – é adotada por muitos teóricos. O filósofo entende que “as mentes são os sistemas de controle dos corpos e para executar suas tarefas determinadas elas precisam coletar, discriminar, armazenar, transformar e processar a informação a respeito das tarefas de controle que executam” (Dennett, 1997, p. 54). A ciência cognitiva – sob a forma de psicologia cognitiva – é uma disciplina que toma o ser humano como um sistema formado por componentes funcionais interconectados, que interagem de forma eficiente e produtiva, estando em plena sintonia com a linguagem funcionalista (cf. Bunnin; Tsui-James, 2002, p. 172).

Não é segredo que os filósofos há tempos vêm tentando compreender a estrutura da mente humana. Esse é o mesmo desígnio das pesquisas em inteligência artificial (IA). A IA é por definição a tentativa de habilitar máquinas (computadores) para a realização de operações que requerem poderes de discernimento que seriam, até onde sabemos, exclusividade de seres humanos (cf. Bunnin; Tsui-James, 2002, p. 173). Os filósofos denominam de IA fraca, a concepção que defende a possibilidade de construir máquinas que agem de forma inteligente. Já a admissão de que tais máquinas, ao realizar ações inteligentes, o fazem de maneira consciente é rotulada de IA forte. Os pesquisadores da área de IA não entram no mérito da discussão filosófica sobre essa dicotomia – assumindo uma postura menos polêmica, afirmam que suas pesquisas intencionam o desenvolvimento de máquinas inteligentes (cf. Gomes, 2010, p. 236).

As pesquisas em IA se iniciaram em meados dos anos 50 e têm revolucionado, através de suas várias áreas, o mundo da tecnologia. As pesquisas realizadas por Warren MacCulloch e Walter Pitts são consideradas por muitos um divisor de águas nos trabalhos que envolvem a IA. Gomes (2010, p. 236 – grifo nosso) continua sua análise acerca da IA argumentando que:

eles se basearam em três fontes: ‘o conhecimento da fisiologia básica e da função dos neurônios do cérebro, uma análise formal da lógica proposicional criada por Russell e Whitehead e a teoria da computação de Turing’. Esses pesquisadores sugeriram um modelo de neurônios artificiais, no qual, cada neurônio era caracterizado por *ligado* ou *desligado*, desse modo, o estado de um neurônio era analisado como, ‘equivalente em termos concretos a uma proposição que definia seu estímulo adequado’.

A primeira edificação de uma visão mais ampla da IA foi realizada por Alan Turing<sup>29</sup>. De acordo com o famoso teste de Turing, se uma máquina apresenta um comportamento indistinguível de um ser humano, não há motivos para negar uma atividade mental para tal entidade. A atividade computacional envolve uma máquina que “recebe, interpreta, processa, armazena, manipula e utiliza informação” (Bunnin; Tsui-James, 2002, p. 173). Ao realizar uma tentativa de programação inteligente, os pesquisadores dessa área estão cientes disso. De acordo com Bunnin e Tsui-James (*Idem*), um problema relativo à IA se apresenta quando:

Dado que a máquina vê isto como *input*, o que ela já precisa saber e o que ela precisa fazer de acordo com isso para ser capaz de (reconhecer, identificar, selecionar, reunir, prever, informar, e assim por diante)? E como, portanto, podemos fazê-la saber isso e levá-la a efetuar essas coisas? Assim, é razoável atribuir tal sucesso da IA a uma confiança autoconsciente do paradigma do processamento da informação.

Os teóricos da IA subscrevem a tese de que a inteligência humana, bem como a cognição, está de algum modo relacionado com o processamento de informação computacional. Esse é o ponto de convergência entre a IA e a teoria da consciência dennettiana.

A inteligência artificial, consoante com a abordagem de Dennett (2006), é uma investigação teórica de cunho analítico, uma abordagem caracterizada por ser “de cima para baixo”. O método aplicado considera o exame, num primeiro momento, de um sistema intencional – uma pessoa (ou uma de suas habilidades, como, por exemplo, fazer uma cadeira) ou um organismo cognitivo – e na sequência subdivide esse sistema em uma economia de subsistemas. Esses subsistemas serão analisados como sistemas completos e serão também fragmentados para análises. Esse processo seguirá até que se chegue no limite em que possam ser tomados como homúnculos e suas especializações – que realizam suas funções em um fluxograma estabelecido por níveis cada vez mais alto de ignorância (cf. Dennett, 2006, p. 180). Alguns filósofos acusam a IA de ser uma petição de princípios, alegando que apresenta um problema incorrigível: oferece apenas soluções mecanicistas, ou seja, não descreve todos os modos possíveis de inteligência, mas tão somente aqueles que são possíveis de realizar, que sejam passíveis de reprodução (cf. Dennett, 2006, p. 167). Todavia, segundo o filósofo, essa acusação se enquadraria bem melhor à psicologia. Em defesa da inteligência artificial, Dennett (2006, p. 167) trata a questão ainda mais a fundo:

---

<sup>29</sup> Delineada em seu artigo de 1950 *Computing Machinery and Intelligency* (cf. Gomes, 2010, p. 236).

estou argumentando que a IA compartilha com a filosofia (em particular, com a epistemologia e a filosofia da mente) o estatuto da investigação mais abstrata dos princípios da psicologia. Mas ela compartilha com a psicologia, de forma distinta da filosofia, uma tática típica de responder às suas questões. Na IA ou na psicologia cognitiva a tentativa típica de responder a questões gerais de cima para baixo consiste em projetar um sistema particular que faz, ou parece fazer, o trabalho relevante, e então considerar quais de seus aspectos são necessários não apenas para um sistema particular de alguém, mas para qualquer sistema como aquele.

A abordagem da inteligência artificial, de acordo com Dennett (2006), entendida como uma análise abstrata da possibilidade de conhecimento e inteligência, solucionou muitos dos problemas gerais que atormentaram filósofos e psicólogos<sup>30</sup>. Acreditava-se que a única descrição possível da complexidade da atividade humana seria a que realizasse representações internas. Além disso, uma representação, ou sistemas de representações, depende de pelo menos um intérprete. Assim, essa abordagem psicológica implicava a ideia de um tipo de homúnculo (cf. Dennett, 2006, p. 178). Com efeito, apesar da impossibilidade de uma psicologia sem homúnculos, essa análise configura-se como uma espécie de *virtus dormitiva*. Sem lograr êxito, Hume tenta solucionar esse problema estabelecendo a noção de que a ideia e a percepção constroem representações para si mesmas. Contudo, apenas com o advento da IA e a noção de estrutura de dados foi possível efetivamente estabelecer representações que se entendem a si mesmas.

Na perspectiva da abordagem da IA, de acordo com Dennett (2006), a inserção do conceito de homúnculo ganha novo entendimento. Utilizando o método de análise “de cima para baixo” cada homúnculo tem uma capacidade de retenção de informação que obedece uma estrutura hierárquica que vai dos mais inteligentes (os que possuem uma linguagem mais expandida) aos mais estúpidos (os que teriam uma linguagem mais limitada) evitando assim a regressão ao infinito (cf. Dennett, 2006, p.180). Sobre esse ponto da teoria, Dennett (2006, p. 181 – nossa tradução) também nos ajuda:

não se consegue completamente nunca uma representação que se autocompreenda (a menos que se recue e veja toda representação no sistema de um ponto de vista privilegiado); mas todos os homúnculos são, em última instância, eliminados. Consegue-se a vantagem do equilíbrio apenas sacrificando a versatilidade e a universalidade dos subsistemas e suas representações; assim, nossos homúnculos não podem ser tão versáteis, nem podem as mensagens que eles enviam e recebem possuir o inteiro sabor da interação linguística humana.

De acordo com Dennett (1991a; 1998), esses agentes são os conectores basilares no campo da inteligência artificial e da ciência da computação em geral. Qualquer um que esteja

---

<sup>30</sup> O denominado problema de Hume é uma discussão levantada sobre a percepção. Ao considerar a ideia de que a percepção visual, por exemplo, se dá pelo processo de representação para um olho interno, teria este que ter também o seu olho interno e este segundo o seu – configurando uma regressão ao infinito (cf. Dennett, 2006, p. 179).

armado com um ceticismo sobre a natureza dos homúnculos, está cego em relação ao alto grau de neutralidade desse conceito, bem como o amplo poder de sua aplicabilidade. Assim, estabelecer um grupo de homúnculos seria irrelevante, se não fosse considerado que, em sua concepção de homúnculos, os aspectos relevantes residem no modo como “interagem, se desenvolvem, formam coalizões ou hierarquias, e assim por diante” (Dennett, 1991a, p. 261 – nossa tradução).

A inteligência artificial, de acordo com Boucher (2020), é um campo que está em pleno desenvolvimento – refletida em pesquisas sobre redes neurais e na elaboração de conceitos de aprendizagem. Entrementes, quando associamos suas descobertas a uma teoria da mente, surgem os tabus. A IA forte, como veremos no capítulo III, é duramente criticada por Searle. Não há unanimidade em relação à viabilidade da *superinteligência artificial*<sup>31</sup> – capaz de realizar as mesmas tarefas que os seres humanos e apresentar os mesmos comportamentos. Alguns defendem que para criá-la será preciso considerarmos os princípios básicos da inteligência humana.

### **1.3 A neurociência e o problema fácil<sup>32</sup>**

Na parte mais anterior dos lobos frontais encontra-se a região mais obscura do nosso cérebro: os lobos pré-frontais. Segundo Ramachandran (2014), lesões nessa região causam estranhos comportamentos. Nela localizamos os intrigantes neurônios-espelho. Sua importância e funcionalidade são fundamentais para entendermos a consciência. Eles podem estar relacionados não só com o aprendizado social, mas também com o processo de transmissão cultural através da imitação – inclusive da linguagem. Esses neurônios são um grupo de células especializadas em leitura mental dentro do cérebro. Aparentemente, eles são a “base fisiológica que faltava para certas habilidades de alto nível que os neurocientistas vinham se sentido há muito tempo desafiados a explicar” (Ramachandran, 2014, p. 183).

O alto grau de desdobramento dos neurônios-espelho, proporcionado pela evolução natural, elevou a cultura a um novo patamar – ela pode ser considerada agora um novo tipo de genoma. Equipados com a cultura, “os seres humanos puderam se adaptar a novos ambientes hostis e descobrir como explorar fontes de alimentos antes inacessíveis ou venenosos em apenas uma ou duas gerações” (Ramachandran, 2014, p. 44). Com efeito, a aquisição da cultura

---

<sup>31</sup> Em termos de desenvolvimento, há três níveis de IA: a já citada *superinteligência artificial*; *Narrow A.I.* (inteligência artificial limitada) – consegue aprender e realizar uma tarefa específica; *General A.I.* (inteligência artificial generalizada) – capaz de realizar mais de uma tarefa específica e de se adaptar para outras tarefas específicas. Até agora só foi possível desenvolver a narrow AI (cf. Boucher, 2020, p. 06-07).

<sup>32</sup> Segundo Chalmers (1996), o exame ontológico da consciência configura o problema difícil, sendo o problema fácil a análise dos aspectos que envolvem o cérebro humano, ou seja, os estudos da neurociência.

acelerou o processo evolutivo da espécie humana, ou seja, evitou a lentidão exigida pela evolução genética que perpassaria centenas ou milhares de gerações.

Consoante Ramachandran (2014), o cérebro possui características neotênicas<sup>33</sup> e apresenta uma relação dinâmica e equilibrada com o mundo e com outros cérebros. Esse sistema extraordinário está, a todo momento, sendo submetido às exigências realizadas por mudanças no ambiente. As investigações sobre os neurônios-espelho, revelam que nossos cérebros estão em plena sintonia com outros, numa dinâmica similar a uma rede social, transformando-se e nutrindo-se mutuamente a todo momento (cf. Ramachandran, 2014, p. 63). Mas como se dá o processo de percepção no sistema nervoso? O que o cérebro realiza com os dados fornecidos pelos olhos? Se partirmos do pressuposto de que o processo visual é uma simples retransmissão estaremos cometendo um erro abissal – já que a imagem é captada de modo invertida se faz necessário um segundo processo, ao longo do nervo ocular para o lobo occipital, para torná-la compatível com o real, mas nesse caso nos conduzirá à chamada falácia do homúnculo – um processo de reinterpretação de imagens realizada por um “homenzinho” dentro de nossas cabeças. Por conseguinte, esse homenzinho teria também o seu homenzinho, e esse último o seu, de tal modo que o processo se configura em uma regressão ao infinito de homúnculos. Em relação a esse problema Ramachandran (2014, p. 74) afirma que:

para compreender a percepção, precisamos primeiro nos livrar da noção de que a imagem do fundo do seu olho é simplesmente retransmitida para nosso cérebro para ser exibida numa tela. Em vez disso, temos de compreender que a partir do momento em que os raios de luz são convertidos em impulsos neurais no fundo de nossos olhos, não faz mais nenhum sentido pensar na informação visual como uma imagem. Temos de pensar, em vez disso, em descrições simbólicas que *representam* as cenas e objetos que compõem a imagem.

Nesse sentido, o cérebro realiza descrições simbólicas. Seu trabalho não se resume à recriação da imagem original. Não obstante, ele estabelece uma representação que engloba as características e aspectos gerais em termos articulados em seus “caracteres” de impulsos nervosos. Esse processamento se dá em um nível que não é alcançado pela consciência (cf. Ramachandran, 2014, p. 75). Salvo a área temporal média (TM) e uma área chamada de V4<sup>34</sup> – que é associada ao reconhecimento – localizada no lobo occipital, a articulação de informações associadas à visão consciente são distribuídas por todo córtex cerebral humano (cf. Ramachandran, 2014, p. 91).

---

<sup>33</sup> Neotenia é o prolongamento de aspectos infantis e juvenis de nossos cérebros que proporciona maior plasticidade durante toda nossa existência (cf. Ramachandran, 2014, p. 63).

<sup>34</sup> A área V4 (fluxo ventral “o que”) se articula com a área da memória para estabelecer o reconhecimento. Já a área temporal média (TM) se articula com áreas do lobo parietal para designar a ação ou comportamento mais adequado (cf. Ramachandran, 2014, p. 92).

Voltando aos neurônios-espelho, este grupo de neurônios está relacionado com as seguintes funções: empatia, dedução de intenção, mímica, fábula e aprendizagem da linguagem. Vale ressaltar que “todas essas atividades requerem adoção do ponto de vista do outro” (Ramachandran, 2014, p. 183-84). Como foram detectados neurônios-espelho no giro supramarginal, podemos inferir que um sistema formado por esse tipo de neurônio nos humanos está relacionado tanto com o processo de interpretação de ações especializadas, quanto à “compreensão de metáforas de ações e, na realidade, em outros aspectos de cognição incorporada” (Ramachandran, 2014, p. 188).

De acordo com Leal-Toledo (2010), os estudos sobre neurônios-espelho apresentam novas implicações relacionadas ao estudo da mente. A psicologia popular postula uma interpretação equivocada do funcionamento do cérebro, ao considerá-lo produtor da mente: o processo cerebral que permite compreender e imitar uma ação que está vinculada à utilização de áreas diferentes do cérebro (cf. Leal-Toledo, 2010, p. 181). Não obstante, as pesquisas da neurociência acerca dos neurônios-espelho indicam que não ocorre nada semelhante. Áreas<sup>35</sup> do cérebro que, até então, eram tomadas como especializadas em funções motoras, demonstraram realizar atividades múltiplas. Esses grupos de neurônios apresentam dois aspectos fundamentais: 1) respondem a atos intransitivos que estabelecem um ambiente propício ao surgimento da linguagem. Gestos e expressões corporais que sugerem transmitir um significado não são associados diretamente a qualquer objeto particular, proporcionando maior poder de compreensão e de imitação dos neurônios-espelho; 2) possuem o atributo de reproduzir, em um alto grau de fidelidade, movimentos observados – respeitando esse caráter peculiar, temos a capacidade de distinguir, por exemplo, os movimentos de uma dança e dos gestos específicos de uma arte marcial. De forma bastante eloquente, Leal-Toledo (2010, p 185) contribui conosco:

É importante notar que o próprio modo como tais neurônios funcionam já nos dá uma excelente indicação da sua importância para a comunicação, pois para um ato comunicativo ter sucesso deve haver uma espécie de paridade. Isso quer dizer que foi devidamente comunicada a mensagem que foi enviada. Sem isso podemos dizer que tal comunicação falhou. Para que isso seja possível parece ser necessário que uma mesma ação seja compreendida de uma forma razoavelmente idêntica em cérebros diferentes.

Ao que parece, os neurônios-espelho são excitados pela compreensão do significado de ações realizadas ou percebidas (cf. Leal-Toledo, 2010, p. 183). Existiria, portanto, uma ligação entre sistemas espelho<sup>36</sup> e a linguagem. Esse entendimento ganhou mais força quando “tal

---

<sup>35</sup> A área motora F5 é um exemplo.

<sup>36</sup> Grupo de neurônios-espelho (cf. Leal-Toledo, 2010, p. 180).

ligação entre neurônios-espelho e a espantosa habilidade humana para a linguagem se tornou ainda mais evidente quando tais neurônios foram descobertos na área de Broca, uma área notoriamente ligada à linguagem” (Leal-Toledo, 2010, p. 184). Grande parte dos neurônios-espelho estão relacionados com atos de alimentação (mastigar, pegar, etc.) e tais ações são bastante similares ao ato de falar. Foram demonstrados em experimentos que, nos humanos, grupos de neurônios-espelho são excitados quando um indivíduo presencia outro (seja um homem, macaco ou cão) mascando um alimento. Não obstante, houve comportamentos assimétricos quando submetido a atos comunicativos sem a emissão de sons, ou seja, diante da visão dos movimentos dos lábios de um homem, um macaco e um cão. Os resultados dos experimentos apontaram para uma elevada excitação desses neurônios quando um ser humano era submetido, uma mais branda foi percebida no macaco e nos cães era praticamente nula (cf. Leal-Toledo, 2010, p. 185).

Esse experimento nos sugere que, em seres humanos, os neurônios-espelho possuem a prerrogativa adicional de proporcionar atos de comunicação, ou seja, o modo pelo qual funciona o processo da comunicação. Sobre esse ponto, Leal-Toledo (2010, p. 186 – grifo nosso) no fornece mais detalhes:

as áreas correspondentes hoje à linguagem (Broca e Wernicke) se situam na região temporal do hemisfério esquerdo, causam certa assimetria do crânio que já começa a ser encontrada nos *Homo habilis*. Por esse motivo, cerca de dois milhões de anos atrás é também a data estimada para o início das pressões seletivas para uma vocalização aumentada [...] que ocasionaram o surgimento do que hoje chamamos de linguagem. Vemos, então, e provavelmente não por coincidência, o nosso principal instrumento no sistema espelho surgindo praticamente juntos.

Nesse sentido, os neurônios-espelho seriam os fundamentos da nossa linguagem, da nossa compreensão e da nossa capacidade de aquisição cultural – através da ação de imitar e de outros recursos de aprendizagem. As pesquisas sobre o comportamento dos neurônios-espelho demonstram que os processos cognitivos – calcados na imitação – são, acredita-se, extremamente distintos dos processos intuitivos nos quais depositamos tanta fé. O cérebro não trabalha em uma dinâmica processual em que primeiro percebe, depois traduz para o corpo e finalmente coordena a ação. Ele, considerando o princípio da economia, procura realizar essa tarefa de modo mais simples, ou seja, não há uma divisão por zona de atribuições. Assim, um mesmo neurônio pode ter múltiplas atribuições – não haveria os neurônios especialistas (cf. Leal-Toledo, 2010, p. 187).

A atividade cerebral, na esteira de Leal-Toledo (2010), que envolve o fazer e o compreender, estão intimamente relacionados – muito mais do que imaginamos – haja vista que um neurônio-espelho é excitado tanto na observação de uma ação específica quanto no agir

efetivo. Esse comportamento é observado também em neurônios relacionados à fala, pois são ativados quando vemos alguém falando, mesmo se não estivermos ouvindo o som, mas não são ativados quando vemos alguém movimentando a boca com outro propósito. Nesse sentido, esses neurônios possuem a atribuição de perceber movimentos relacionados exclusivamente a atos de comunicação verbal – já que “é possível que algumas das nossas funções cerebrais, parte do que normalmente gostamos de chamar de mente, possam ser entendidas por meio da ativação de neurônios-espelho” (Leal-Toledo, 2010, p. 187). A psicologia popular trata o pensar e o fazer como categorias distintas. Realizar uma corrida e pensar em correr seriam duas coisas extremamente diferentes. Todavia, não é o que indicam as pesquisas sobre neurônios-espelho. Esses estudos nos fornecem uma possibilidade de abordar o problema mente/cérebro de modo inovador, pois é possível que a “diferença entre realizar um movimento físico e pensar neste movimento pode ser apenas que no segundo caso uma área inibitória do nosso cérebro está agindo” (*Ibidem*), e, nesse caso, nosso cérebro pode despender maior labor para inibir uma ação do que para realizá-la. Por conseguinte, antes da evolução presentear a espécie humana com uma área para inibir ações observadas, o pensamento seria impossível.

De acordo com Leal-Toledo (2010), esse entendimento se choca não só com as concepções da psicologia popular, mas também com algumas teorias da mente calcadas na noção de representação mental, ou interna. Esse tipo de abordagem é denominado de representacionalismo e é uma das características fundamentais das teorias dualistas – além de ser percebida também em algumas teses fisicalistas.

Portanto, faz mais sentido admitir que o processo é muito mais diversificado. As pesquisas apontam que ao invés de conceber um centro decisório como realizador ou não da ativação das áreas motoras, o que ocorre na verdade é que o centro inibitório tem o atributo de coibir ou não as áreas motoras. Sobre esse ponto Leal-Toledo (2010, p. 190) aduz que:

a área inibidora não é um centro decisório, pois um centro decisório deve decidir se queremos ou não fazer um determinado movimento. Só quando ele decidir que quer fazer o movimento é que ele manda o sinal para as áreas motoras que aguardam passivamente. Já o centro inibitório apenas reprime um movimento que já está em vias de ser realizado automaticamente.

Esta nova abordagem sobre as operações realizadas pelo nosso cérebro é de caráter contra-intuitiva. O dualismo tem, em sua defesa, o argumento de que temos o acesso direto aos estados mentais. Só que, não há motivos para conceber o funcionamento do cérebro em uma perspectiva intuitiva. Mais uma vez, o oposto nos parece muito mais palatável, ou seja, é muito mais complicado, talvez impossível, examinar os processos subjacentes à própria intuição, e

muito mais admissível que o nosso cérebro (e portanto nossa mente) trabalhe de maneira que só pode ser entendido através de uma análise contra-intuitiva (cf. Dennett, 1991a, p. 62).

Com efeito, de acordo com Leal-Toledo (2010), as grandes teorias científicas são caracterizadas por previsões contra-intuitivas. Essas previsões teóricas são de cunho evolutivo e, portanto, se adaptam a um ambiente específico. A matemática e a física, por exemplo, estavam ambientadas na concepção euclidiana e newtoniana, respectivamente. Destarte, se considerássemos exclusivamente abordagens que estivessem em consonância com nossas intuições, seria inconcebível adaptá-las a um ambiente einsteiniano ou quântico, pois essas descrições do mundo físico são extremamente contra-intuitivas (cf. Leal-Toledo, 2010, p. 191). Quando objetos se aproximam da velocidade da luz, a variação do tempo é mais lenta e os objetos se comprimem na proporção em que suas massas aumentam; as partículas não apresentam estado definido, segundo a mecânica quântica, etc. Essa característica contra-intuitiva também é percebida na biologia populacional, contrariando nossas intuições sobre a origem das espécies de vida.

Consoante Leal-Toledo (2010), para manter a vida – dentro da cadeia da evolução – a espécie humana (o seu cérebro) não necessitou compreender a sua estrutura funcional interna. O conhecimento dos mecanismos que proporcionam a identificação de um alimento ou de um predador são irrelevantes. Nos primórdios da espécie humana, como ainda hoje, não necessitamos de uma descrição rigorosa do funcionamento de nosso cérebro para que ele execute bem suas funções. O cérebro nunca precisou relatar a estrutura subjacente de seu funcionamento, os resultados desses processos laborativos é o aspecto intuitivo que se nos apresenta. Os antigos argumentos, nossas descrições mais primitivas sobre o funcionamento da coisa-mente, que são calcadas no testemunho direto de nossas intuições, não podem fundamentar a sua verdade; nossas intuições mais essenciais são insuficientes para descrevê-la; as abordagens da coisa-mente em primeira pessoa podem estar equivocadas – na verdade é muito provável que estejam, se considerarmos que no seu surgimento as necessidades dos seres humanos eram outras (cf. Leal-Toledo, 2010, p. 191). Sobre o erro do representacionalismo, o comentário de Leal-Toledo (2010, p. 192) é esclarecedor:

A noção de representacionalismo e o modo como ela é usada nas neurociências pelos “materialistas cartesianos” é um excelente exemplo de como ambas as noções se confundem: busca-se no cérebro o local para aquilo que acreditamos existir por causa das nossas intuições sobre o funcionamento de nossa mente.

O estudo sobre o funcionamento do nosso cérebro nos ajudará a delinear um novo modo de entendimento da estrutura que dá origem aos processos mentais. As pesquisas sobre

neurônios-espelho contribuem para esse empreendimento, haja vista que exigem que as nossas convicções sobre a coisa-mente passem por uma revisão.

Destarte, a explicação das intuições deve ser perseguida, mas o fato de tentarmos explicá-la não nos dá o direito de excluir categoricamente uma abordagem contra-intuitiva, ou seja, ela não precisa ser uma descrição estritamente intuitiva – na verdade seria mais lógico que seja iniciada por uma abordagem contra-intuitiva. Devemos considerar, portanto, que a explicação da coisa-mente com maior plausibilidade será oriunda de uma abordagem contra-intuitiva, pois “nossas intuições são aquilo que acreditamos que não precisa ser explicado” (Leal-Toledo, 2010, p. 192).

A possibilidade de uma teoria bem-sucedida, ao que parece, está sendo proporcionada pelas revelações das pesquisas sobre os neurônios-espelho. Portanto, a cisão entre o comportamento e pensamento – descritos pelos críticos do behaviorismo – deve ser revista. Entrementes, não deva ser o objetivo maior a retomada integral da teoria do comportamento, devemos considerar alguns de seus aspectos fulcrais, mais precisamente aqueles que reconectam a mente e o comportamento (cf. Leal-Toledo, 2010, p. 192).

De acordo com Ramachandran (2014), essa tarefa requer um estudo sobre as áreas do cérebro envolvidas com o desenvolvimento da linguagem. Segundo ele, possivelmente a especialidade da área de Broca não se resume apenas à sintaxe – ou a estruturas da linguagem natural. Ela pode estar associada também a linguagens arbitrárias que possuem regras formais (tais como programação de computador ou álgebra). A complexidade da análise da estrutura cerebral, concernente à linguagem, é um desafio para os cientistas:

A linguagem humana parece tão complexa, multidimensional e ricamente evocativa que somos tentados a pensar que quase todo o cérebro deve estar envolvido, ou pelo menos grande parte dele. Afinal, mesmo a pronúncia de uma única palavra como “rosa” evoca toda uma miríade de associações e emoções: a primeira rosa que você ganhou na vida, o perfume, os jardins de rosas que lhe prometeram, face e lábios rosados, espinhos, óculos cor-de-rosa e assim por diante. Não implicaria isso que muitas regiões do cérebro distante umas das outras devem cooperar para gerar o conceito de uma rosa? Sem dúvida a palavra é apenas a maçaneta, ou o foco, em torno da qual gira um halo de associações, significados e lembranças (Ramachandran, 2014, p. 206).

No entanto, ocorre exatamente o contrário, ou seja, os estudos de pacientes vítimas de afasias apontam que os cérebros possuem áreas especializadas em linguagem<sup>37</sup> – embora devamos concebê-lo como um grande sistema interconectado (Ramachandran, 2014, p. 206).

---

<sup>37</sup> De acordo com Ramachandran (2014), a área de Wernicke está relacionada ao aspecto semântico (significado/pensamento), enquanto que a área de Broca está associada à sintaxe (linguagem).

A linguagem humana é uma protolinguagem<sup>38</sup> com um vasto vocabulário. Ao atingir oito anos de idade, por exemplo, uma criança possui aproximadamente seiscentas palavras introduzidas em seu vocabulário (a memória tem papel fundamental nesse processo). Possuímos palavras cuja significação está vinculada ao contexto da linguagem. Podemos também utilizá-las para designar coisas ou eventos que estão ausentes, ou relacionados com o tempo – passado e futuro – ou ainda em uma realidade hipotética. O uso de metáforas e analogias – apesar de permear a alusiva fronteira entre pensamento e linguagem – é uma exclusividade, até onde se sabe, dos seres humanos. Somos detentores de uma sintaxe flexível, recursiva (cf. Ramachandran, 2014, p. 211).

Há duas áreas de pesquisa, de acordo com Ramachandran (2014), que comumente seduz gênios e loucos, a saber: a consciência e a linguagem. Ao abordar o tema, alguns neurocientistas cognitivos entendem que os linguistas erram ao dar ênfase aos aspectos estruturais da linguagem. Eles não consideram examinar como esses sistemas interagem com outros processos cognitivos. Essa digressão é comentada por Ramachandran, (2014, p. 212):

Eles professam interesse unicamente pelas regras fundamentais para os circuitos gramaticais do cérebro, não pelo modo como os circuitos realmente funcionam. Esse foco estreito remove o incentivo para investigar como esse mecanismo interage com outras capacidades mentais como semântica (que linguistas ortodoxos nem sequer consideram como um aspecto da linguagem), ou para formular questões evolucionárias sobre como ele pode ter evoluído a partir de estruturas cerebrais preexistentes.

Tanto Wallece quanto Chomsky, segundo Ramachandran (2014), atribuem o surgimento da linguagem a um fator miraculoso. No entanto, os estudos neurocientíficos apontam para o entendimento de que determinadas áreas do cérebro foram exaptadas<sup>39</sup>. Há um desconforto na versão de que o pensamento foi desenvolvido primeiro e proporcionou, logo em seguida, o surgimento da linguagem. Essa postura teórica não resolve o problema, mas apenas o procrastina, pois não sabemos ainda o que é o pensamento, ou seja, devemos ser prudentes e evitar tentar “explicar um mistério por outro mistério” (Ramachandran, 2014, p. 216).

---

<sup>38</sup> Conjunto de características que tornam a linguagem humana única, radicalmente diferente de outros tipos de comunicação que vemos em outros animais (cf. Ramachandran, 2014, p. 210).

<sup>39</sup> A exaptação se dá quando um mecanismo, que foi originalmente desenvolvido pela seleção natural para um objetivo específico, abre precedente para o surgimento de algo que possui uma finalidade extremamente distinta. Nesse caso, áreas do cérebro que tinham outras finalidades possibilitaram o surgimento da protolinguagem (cf. Ramachandran, 2014, p. 215).

A abordagem sincinesética<sup>40</sup> se oferece para preencher essa lacuna. Essa tese defende a ideia de que é possível haver uma correspondência similar à observada no efeito bouba-kiki<sup>41</sup> em um nível mais amplo, uma tradução incorporada, ou seja, uma conexão “entre mapas visuais e auditivos, por um lado, e os mapas , motores na área de Broca por outro” (Ramachandran, 2014, p. 222). Com efeito, a linguagem não se reduz às palavras, devemos considerar como se articulam seus aspectos semânticos e sintáticos. Sabemos que essas funções são relativamente autônomas – a afasia de Broca e de Wernicke corrobora essa hipótese: uma pessoa que sofre de afasia de Wernicke, por exemplo, é capaz de desenvolver sentenças complexas e enunciá-las com fluidez numa estrutura gramatical de alto nível, mas sem apresentar um domínio do significado – nos intrigando com o fato de que a área de Broca pode executar seu trabalho com excelência, numa economia de manipulação de palavras e regras gramaticais padrão nos parâmetros de um computador (cf. Ramachandran, 2014, p. 227).

As pesquisas apontam que o desenvolvimento do lobo parietal inferior esquerdo, que pode ter acelerado a protolinguagem e sua subsequente divisão, e ter “desempenhado um papel decisivo na emergência de funções exclusivas humanas” (Ramachandran, 2014, p. 229), nelas incluída um alto nível de abstração. O exame sobre os sintomas da apraxia<sup>42</sup> fundamenta essa hipótese.

As descobertas da neurociência não estão muito distantes do pensamento dennettiano, haja vista que há nele uma defesa que a consciência é uma leitura interna de um padrão de comportamento. Dennett (1991a, p. 171-72 – nossa tradução) nos traz um panorama que facilita mais esse entendimento:

Até agora estamos tratando a própria consciência como uma espécie de caixa preta. Tomamos seu “comportamento” (= fenomenologia) como “dado” e nos perguntamos que tipo de mecanismos ocultos no cérebro poderiam explicá-lo. Agora vamos inverter a estratégia e pensar sobre a evolução dos mecanismos cerebrais para fazer isso e aquilo, e ver se alguma coisa que surge nos dá um mecanismo plausível para explicar alguns dos “comportamentos” intrigantes de nossos cérebros conscientes.

---

<sup>40</sup> A sincinesia é um termo cunhado por Ramachandran (2014) para designar sua teoria da evolução da linguagem, segundo a qual considera que não só a linguagem, mas uma grande parte das características exclusivamente humanas (como pensamento metafórico e a abstração) são oriundas de um gradativo cruzamento de áreas cerebrais, ou seja, essa hipótese defende a ideia de que “a linguagem e muitos aspectos do pensamento abstrato evoluíram por meio de exaptações cuja combinações fortuitas produziram novas soluções” (Ramachandran, 2014, p. 223), se distanciando da ideia de que a linguagem evoluiu do pensamento e de que se desenvolveu como uma estrutura especializada em comunicação.

<sup>41</sup> O efeito bouba-kiki é a “correspondência não arbitrária e a ativação entre mapas cerebrais para visão e sons” (Ramachandran, 2014, p. 222).

<sup>42</sup> Incapacidade de realizar movimentos motores

Consideremos uma analogia entre uma aranha e um ser humano. A aranha, apesar de saber em algum grau a lei de Hooke<sup>43</sup>, não pode aplicar seu “conhecimento” em outras finalidades, pois lhe falta a capacidade de representação abstrata. Por outro lado, o ser humano tem a capacidade de aplicar a lei de Hooke com outros objetivos, ou seja, uma aplicação irrestrita e flexível. Temos a capacidade de representar essa lei – garantida pelo domínio da linguagem (cf. Ramachandran, 2014, p. 236). O gênero humano, através de nosso cérebro, tem a capacidade de criar representações, símbolos arbitrários relativamente semelhantes às palavras, e realizar manipulações para promover uma realidade virtual com o intuito de solucionar um problema ou estabelecer um comportamento – outros animais na mesma condição “*offline*”, um chimpanzé por exemplo, falharia nessa tarefa (cf. Ramachandran, 2014, p. 239).

A realização de exames de pacientes com doenças neurológicas indica que o “eu” (*self*) não é uma substância única que a nossa introspecção nos apresenta. Não obstante, as implicações dos experimentos colidam com nossas intuições mais cristalizadas referentes a nós mesmos, devemos nos ater aos fatos, na medida em que as pesquisas neurológicas nos revelam “que o *self* consiste de muitos componentes, e a noção de um *self* unitário pode ser uma ilusão” (cf. Ramachandran, 2014, p. 312 – grifo nosso).

De acordo com Dennett (1997), a percepção do tempo camufla os padrões dos projetos intencionais, ou seja, há uma interferência em nosso julgamento por parte do tempo, de tal modo que atribuímos mentes aos entes observando a velocidade de seu comportamento (cf. Dennett, 1997, p. 49) – isso configura um chauvinismo de escala temporal. Nesse sentido, um ente para ser caracterizado como possuidor de uma mente deve responder a seu ambiente considerando sua necessidade temporal, sem com isso estabelecer uma introspecção. Sem embargo, para os fisicalistas incautos, a atividade periférica no sistema nervoso é pura e simplesmente sensibilidade – o que abre um precedente para o apontamento de um centro de produção da sciência (cf. Dennett, 1997, p. 57). O autor ainda traça uma importante linha tangencial entre os fisicalistas funcionalistas que defendem a tese de que o cérebro é o centro da atividade funcional do corpo:

Ao abraçar este modo padrão de pensar, ignoramos uma alternativa importante: visualizar o cérebro (e portanto a mente) como um órgão entre muitos, um usurpador relativamente recente do controle, cujas funções não podem ser apropriadamente

---

<sup>43</sup> De acordo com a lei de Hooke, ao aplicar uma sobre uma mola proporcionamos uma deformação – o que a leva a produzir uma força contrária à força externa denominada força elástica. Por definição, essa lei pode ser descrita como uma equação matemática ( $F_{el} = k \cdot x$ , onde  $F_{el}$  – força elástica,  $k$  – constante elástica e  $x$  – deformação da mola) que permite verificar o sentido e calcular o módulo da força elástica que é realizada por molas, quando são submetidas a uma ação de compressão ou esticá-las.

entendidas até que, de algum modo, o vejamos não como o chefe, mas sim como um serviçal, trabalhando para implementar os interesses do corpo que o abriga e o alimenta e dá significado às suas atividades (Dennett, 1997, p. 61).

Em termos concretos, isso significa que o cérebro não deve ser assimilado como um órgão superior, pois um exame etológico da evolução indica que há uma interação entre o sistema nervoso e os demais sistemas – o cérebro calca suas atividades nas informações dos vários sistemas angariados no decorrer da evolução (cf. Dennett, 1997, 69). O processo de exaptação, portanto, seria a égide desse entendimento.

Com efeito, Dennett (1997) está convicto que grande parte dos animais podem ser submetidos ao processo de aprendizagem. Eles são capazes de “replanejar” seu comportamento, ou seja, canalizá-los para novas especificidades, que são estabelecidas em processos de caráter contínuo e de longo prazo, de remodelação pelo ambiente. Em virtude disso, temos “agora bons modelos, com vários graus de realismo e detalhe, de como tal processo de condicionamento ou treinamento pode ser, de modo não milagroso, realizado em uma rede de células nervosas” (Dennett, 1997, 69). De forma bastante eloquente, o filósofo afirma que “a sciência aparece com todos os graus imagináveis ou intensidades, desde o mais simples e mais *robótico* até o mais sofisticadamente sensível, *hiper-reativo humano*” (Dennett, 1997, p. 78 – grifo nosso). Ao comentar a prática da autoestimulação cognitiva, Dennett (1991a, p. 196 – nossa tradução) afirma:

Tudo o que deve ser o caso para que essa prática tenha essa utilidade é que as relações de acesso preexistentes no cérebro de um indivíduo sejam menos do que ideais. Suponha, em outras palavras, que, embora a informação certa para algum propósito já esteja no cérebro, elas estão nas mãos do especialista errado; o subsistema no cérebro que precisa das informações não pode obtê-las diretamente do especialista – porque a evolução simplesmente não chegou a fornecer um “fio” assim. Provocando o especialista a “transmitir” as informações para o meio ambiente, no entanto, e depois confiar em um par de orelhas existentes (e um sistema auditivo) para buscá-lo, seria uma maneira de construir um “fio virtual” entre os subsistemas relevantes.

Consoante Dennett (1991b), boa parte dos elementos de representação são inatos – devido a uma gama de possibilidades proporcionada pela evolução –, mas há um quinhão que deve ser aprendido. Algumas categorias vitais são inseridas numa programação natal e outras devemos angariar em nossa relação com o ambiente. Como isso é possível? Em termos concretos, é a plasticidade do cérebro que permite essa façanha. No entanto, convém que no ambiente haja algo passível de aprendizagem “que já seja produto de um processo de *design* anterior, para que cada um de nós não tenha que reinventar a roda” (Dennett, 1991a, p. 192 – grifo nosso). Na evolução cultural, encontramos um novo meio de evolução associado à plasticidade fenotípica, análoga àquela apresentada na variação genética. A utilização de nossa plasticidade não se resume à simples aprendizagem – procuramos aprimorar e ampliar esse

aprendizado, ou seja, aplicamos um processo evolutivo também na aprendizagem. Com esse intento, disponibilizamos os resultados dessa aprendizagem para a geração seguinte. De algum modo, realizamos *downloads* de hábitos criados e lapidados em cérebros mais ou menos estruturados (cf. Dennett, 1991a, p. 193). Dennett (1991a, p. 182 – nossa tradução) explica melhor esse ponto:

O surgimento da plasticidade nos sistemas nervosos ocorreu ao mesmo tempo (aproximadamente) que os desenvolvimentos que já esboçamos, e forneceu dois novos meios nos quais a evolução poderia ocorrer, em uma velocidade muito maior do que a evolução genética sem ajuda por meio de mutação genética e seleção natural. Uma vez que algumas das complexidades da consciência humana são os resultados de desenvolvimentos que ocorreram e continuam a ocorrer nessas novas mídias, precisamos de uma visão clara, embora elementar, de suas relações entre si e com o processo genético subjacente à evolução.

De acordo com Dennett (1991b), quando examinamos alguns danos mecânicos, intervenções temporárias realizadas por experimentos ou patologias no cérebro, temos uma noção de como é organizada essa estrutura. Embora não tenham sido realizados testes empíricos adequados, tais fenômenos indicam que a teoria dennettiana parece estar mais próxima da verdade (cf. Dennett, 1991a, p. 251). O filósofo não nega que o modelo burocrático tenha a capacidade de explicar essas patologias. Todavia, essas falhas não parecem estar intimamente relacionadas com um sistema burocrático.

#### **1.4. Experimentos de pensamento: bombas de intuições ou muletas de lança?**

Na história da produção de conhecimento, não raro nos deparamos com historietas que auxiliam filósofos e cientistas na abordagem de determinados problemas, um recurso metodológico amplamente utilizado nessas áreas de conhecimento, a saber: os experimentos de pensamento (EPs). Em termos concretos, isso significa que algumas “alegorias e paradoxos são dispositivos da imaginação comumente usados para investigar a natureza das coisas. A simples citação de alguns deles já nos dá a dimensão de sua enorme influência à ciência” (Pereira, 2015, p. 184). No entanto, a eficácia da aplicação dos experimentos de pensamento nas ciências – e talvez, aparentemente, um pouco menos na filosofia – é questionada por parte de alguns teóricos que assumem um empirismo radical.

De acordo com Pereira (2015), três questões são basilares para refletir sobre a relevância dos experimentos de pensamento: o que são? Os experimentos de pensamento podem fornecer uma fonte de conhecimento do mundo natural? De onde viriam esses conhecimentos? Segundo a autora, a função dos EPs é não só ajudar na percepção de contradições de seu argumento inicial, mas também proporcionar uma recontextualização dos dados empíricos antigos.

Em consonância com Cooper (2005), na busca por situar os EPs dentro das diferentes áreas onde emergem, existem dois grupos de filósofos que abordam o exame dos experimentos de pensamento de forma categoricamente distintos: o primeiro assume uma postura cautelosamente científica, ou seja, como não estão certos de que os experimentos de pensamento são iguais tanto na ciência quanto na filosofia se restringem a uma abordagem exclusivamente empirista (cf. Cooper, 2005, p. 329). O segundo admite que há uma impossibilidade de estabelecer uma diferença crucial entre ciência e filosofia no que tange à aplicação dos experimentos de pensamento, pois, por mais das vezes, o que se nota é uma fusão desses campos de investigação ou apagamento das fronteiras teóricas entre essas áreas. De modo ainda mais claro, Cooper (*Ibidem*) nos ajuda:

O trabalho dos filósofos da mente e da linguagem, empiricamente inclinados, são muitas vezes indistinguíveis do trabalho em psicologia teórica ou linguística. O mesmo vale para filósofos da física e físicos teóricos, e para teóricos de jogos, economistas e biólogos teóricos evolucionistas. Em muitos casos, trabalhos filosóficos e científicos só podem ser distinguidos com base nos periódicos em que são publicados. Além disso, por sua natureza necessariamente não empírica, o trabalho envolvendo experimentos mentais é particularmente provável de cair na fronteira entre a filosofia e a ciência.

Há também visões pessimistas, em relação aos experimentos de pensamento, mais radicais. São acusados de suprimir informações relevantes e de um excesso de imaginação (o que dá margem para intuições ambíguas, ou seja, o que uma pessoa acha intuitivamente certo, outra irá considerar obviamente falso). Além disso, segundo essa análise, os EPs promovem uma cisão com a realidade (cf. Brown, 1991, p. 29).

No entanto, segundo Brown (1991), os EPs no campo da filosofia não seriam ilegítimos se estivessem em consonância com as leis da natureza tais como as conhecemos. Nesse sentido, podemos admitir que “imaginar pessoas se dividindo como ameba, certamente viola as leis da biologia. Mas, por outro lado, imaginar-me como um cérebro em uma cuba não” (Brown, 1991, p. 30 – nossa tradução).

Além disso, para que EPs sejam considerados, os detalhes da situação imaginada devem ser de conhecimento geral, ou seja, devem conter elementos familiares à comunidade. Esses tipos de experimentos, no mais das vezes, ajudam a descobrir leis da natureza. De sorte que seria falacioso o argumento que defende a ideia de que um EP seria ilegítimo por violar as leis da natureza. Nesse sentido, ao “estipular as leis com antecedência e exigir experimentos mentais que não as viole seria simplesmente comprometer seu uso como ferramenta poderosa para a investigação da natureza” (Brown, 1991, p. 31 – nossa tradução).

De acordo com Brown (1991), os experimentos mentais são abundantes na física e na filosofia. Entretanto, são raros em outras ciências – na química por exemplo. O porquê dessa

excentricidade, segundo o autor, pode estar relacionado com “a personalidade ou o estilo científico de quem faz física ou filosofia” (Brown, 1991, p. 32), ou quem sabe um acidente histórico, ou seja, a tradição de realizar experimentos de pensamento se desenvolveu apenas em alguns poucos campos (cf. Brown, 1991, p. 32-33). Além disso, no exame da estrutura dos EPs, apesar de assumirem múltiplas facetas, é possível estabelecer alguns padrões. Esses, segundo o autor, podem ser classificados em dois grupos gerais, a saber: experimentos de pensamento destrutivos (EPDs) e experimentos de pensamento construtivos (EPCs). Os experimentos de pensamento destrutivos intencionam refutar uma teoria, e podem ser caracterizados como “uma pitoresca *reduction ad absurdum*; destrói ou pelo menos apresenta alguns sérios problemas de uma teoria, geralmente aponta uma lacuna em seu quadro geral” (Brown, 1991, p. 34). Esse tipo de experimento invariavelmente leva o autor da tese que se pretende refutar a reformular seu argumento. No entanto, sua eficácia muitas vezes é comprometida por deixar de fora alguns aspectos da teoria em questão que são fundamentais para sua compreensão.

Portanto, a classificação realizada para delinear o EP construtivo pede maior rigor, ou seja, a emergência de três grupos específicos que dão origem aos seguintes conceitos: primeiramente, teríamos os experimentos de pensamento construtivos mediativos – visam sempre facilitar o entendimento de uma conclusão extraída de uma teoria específica e bem articulada.

Segundo Brown (1991), a taxonomia aplicada aos experimentos de pensamento construtivos revela uma série de subnotificações, ou seja, uma subclassificação que não comprometem o seu entendimento. Portanto, a classificação realizada para delinear o EP construtivo pede maior rigor, expresso na emergência de três grupos específicos que dão origem aos seguintes conceitos: primeiramente, teríamos os experimentos de pensamento construtivos *mediativos* – visam sempre facilitar o entendimento de uma conclusão extraída de uma teoria específica e bem articulada. Por ora, vejamos o que Brown (1991, p. 37 – nossa tradução) diz a esse respeito:

um experimento de pensamento mediativo pode ilustrar alguns casos contrários, aspectos altamente contra-intuitivos da teoria, portanto, fazendo com que pareça mais palatável; ou pode atuar como um diagrama em uma prova geométrica na medida em que nos ajuda a entender a derivação formal que pode ter sido descoberta essencialmente na prova formal.

Assim, por serem apresentadas sempre após a teoria, o desígnio desse tipo de experimento de pensamento é demonstrar que os aspectos considerados impossíveis de uma teoria são, na verdade, logicamente coerentes (cf. Brown, 1991, p. 37).

Uma segunda subclassificação dos EPCs são os *conjecturais*. Esse tipo procura estabelecer o entendimento da possibilidade de realização de determinado fenômeno. Caracteriza-se por preceder a teoria – o aspecto conclusivo da argumentação (cf. Brown, 1991, p. 40). Nesse sentido, somos levados a conjecturar uma hipótese a partir do experimento apresentado.

Consoante ao entendimento de Brown (1991), a terceira subclassificação pode ser denominada de experimento de pensamento *direto*. São similares aos EPs mediativos – que se iniciam com fenômenos não problemáticos –, mas se distinguem por demonstrarem que “a relação lógica entre o princípio geral vago e a conclusão simplesmente não existe” (Brown, 1991, p. 43).

Examinemos agora, ainda na esteira da taxonomia dos experimentos de pensamento edificado por Brown (1991), os experimentos de pensamento *platônicos*. Esse tipo de EP é percebido, segundo o filósofo, nos dois tipos gerais de experimentos – construtivos diretos e destrutivos. Há casos, porém, em que a aplicação dessa classe de experimento enseja refutar a antiga teoria e concomitantemente corroborar o novo argumento – o autor exemplifica esse tipo de experimento utilizando o relato de Galileu Galilei (1564-1642) sobre a queda livre. Destarte, os EPs platônicos nos permitem pensar sobre objetos abstratos e relações universais.

A taxonomia dos experimentos de pensamento edificada por Robert Brown nos ajuda a perceber sua relevância como método. No entanto, apesar dos EPs ajudarem a proporcionar um conhecimento sobre o mundo natural, se forem mal elaborados podem culminar em falhas e mal-entendidos. Vale ressaltar que, ainda que os experimentos mentais, geralmente, envolvam raciocínios contrafactuais (não experimentais) eles são fortemente sensíveis ao contexto. Para serem bem-sucedidos devem estar com os pés enfiados na lógica, se realmente pretendem corroborar ou refutar uma teoria debatida.

Deferindo uma crítica indireta à taxonomia browniana – já que seu desígnio maior é examinar a posição epistêmica e metafísica de Brown –, John Norton (1996) irá rebaixar o *status* dos experimentos de pensamento a meros argumentos. Em detrimento do entendimento de que EPs revelam conhecimentos ou contradições epistemológicas, o autor, ao delinear a *tese da reconstrução*<sup>44</sup>, demonstra que esse recurso não promove milagres epistêmicos, ou seja, “na medida em que podem nos contar algo sobre nosso mundo, eles o fazem usando nossos recursos epistêmicos padrão: experiências comuns e as inferências que extraímos delas” (Norton, 1996,

---

<sup>44</sup> Os “experimentos podem ser reconstruídos como argumentos baseados em suposições tácitas ou explícitas. A crença na conclusão do resultado do experimento mental é justificada apenas na medida em que o argumento reconstruído pode justificar a conclusão” (Norton, 1996, p. 338 – nossa tradução).

p. 334 – nossa tradução). Nessa perspectiva, o autor concebe os experimentos de pensamento como simples argumentos pitorescos, desnudos da imagem exagerada de notoriedade epistêmica.

De acordo com Norton (1996), todo e qualquer experimento de pensamento deve ser de caráter ordinário. Nesse sentido, seu pensamento é oriundo de uma filosofia conservadora de cunho empirista. Para ele, um filósofo sério<sup>45</sup> deve se abster de qualquer sugestão que o leve a abandonar um dos principais fundamentos das experiências nas ciências, as experiências reais, para adotar as fantasias da imaginação travestidas de argumentos lógicos (cf. Norton, 1996, p. 335). Norton é categórico. Ele está convencido de que EPs não utilizam dados experimentais novos, porquanto se limitam apenas a reorganização ou generalização daquilo que já conhecemos (ou experienciamos) sobre o real. Sobre a sua posição, Norton (1996, p. 335 – nossa tradução) nos oferece mais detalhes:

Esse conhecimento prévio repousa eventualmente sobre a experiência e só pode entrar em experimentos mentais como suposições, tácitas ou explícitas. Portanto, experimentos mentais são dispositivos que reorganizam ou generalizam essas suposições para produzir o resultado do experimento mental. Ou seja, esses dispositivos são argumentos que nos levam de nossas suposições a uma conclusão, o resultado do experimento mental. Na medida em que o dispositivo meramente reorganiza, é um argumento dedutivo; na medida em que generaliza no sentido mais amplo, é um argumento indutivo.

A conclusão é admitida ao passo que nossas predições são verdadeiras e as inferências válidas – isso não garante, no entanto, que os EPs possuem a qualidade de proporcionar inferências dedutivas ou indutivas perfeitas<sup>46</sup>. Ao examinarmos os EPs como recursos epistemológicos os concebemos como argumentos<sup>47</sup>. Para Norton, é isso que eles são. Nessa perspectiva, um “bom experimento mental é um bom argumento; um experimento de pensamento ruim é um argumento ruim” (*Ibidem*).

Segundo Norton (1996), os experimentos de pensamento se caracterizam como argumentos que postulam estados de coisas hipotéticos ou contrafactuais e invocam particularidades irrelevantes para a generalidade da conclusão. De acordo com o autor, não vale a pena produzir uma taxonomia dos experimentos de pensamento e critica a realizada por Brown, pois ele entende que “os detalhes desta taxonomia têm pouca relevância para os epistemólogos” (Norton, 1996, p. 338 – nossa tradução).

---

<sup>45</sup> Os filósofos das ciências empíricas, segundo Norton.

<sup>46</sup> Os experimentos mentais podem ser malfeitos, assim como os argumentos.

<sup>47</sup> A Tese da eliminação (elimination thesis) sugere que é perfeitamente plausível a substituição de um experimento de pensamento por um argumento.

Na abordagem de Brown (1991), estamos diante de uma epistemologia não empirista cujo caráter especial *a priori* é o fundamento dos experimentos de pensamento platônico. Seguindo o prisma do platonismo matemático, esses EPs são apresentados como ferramentas para detectar as leis da natureza, do mesmo modo como percebemos os objetos em nosso cotidiano. Porquanto, em consonância com a análise da taxonomia browniana constata-se que “essas leis têm uma existência real dentro de um mundo platônico, assim como os objetos cotidianos têm uma existência real no mundo da experiência cotidiana” (*Ibidem*). Norton oferece esse entendimento de forma clara e vigorosa ao citar Brown:

Este conhecimento prévio é adquirido por um tipo de percepção das leis da natureza que são, argumenta-se, interpretadas de forma realista. Assim como a mente matemática pode compreender (alguns) conjuntos abstratos, a mente científica pode compreender (algumas) entidades abstratas que são as leis da natureza (*Idem*).

Em termos concretos, isso significa que Brown obtém respaldo teórico na falta de qualquer inferência mediadora possível no surgimento da conclusão do experimento de pensamento. Essa fé na ausência da aparente inferência mediadora leva Brown a rotular a crítica nortoniana de antitética em relação a sua (cf. Norton, 1996, p. 338).

Norton (1996) se recusa a acreditar que em um EP platônico realiza-se o milagre de promover percepções que destoem daquelas extraídas da experiência mundana comum. Demonstrando que as mesmas conclusões podem ocorrer se executarmos os meios ordinários “bem podemos começar a nos perguntar se precisamos dos mistérios da percepção platônica para explicar experiências de pensamento” (Norton, 1996, p. 340). Como já foi dito, o autor demonstra que é possível transformar o experimento de pensamento platônico em argumento aplicando a tese de reconstrução. A respeito disso Norton (1996, p. 357 – nossa tradução) também nos ajuda:

O sucesso da tese de reconstrução mostra que a visão argumentativa é viável. É, insisto, também a visão mais parcimoniosa dos experimentos mentais. Uma vez que a argumentação é incontrovertidamente uma parte de muitos experimentos de pensamento, qualquer explicação razoável de experimentos de pensamento deve dar algum lugar a ela. Dado que a argumentação também é suficiente para fornecer uma explicação completa da epistemologia dos experimentos de pensamento, por que deveríamos considerar recorrer a qualquer maquinaria epistêmica adicional? A navalha de Ockham nos instrui a não fazer isso. O mago materializa uma pomba. Sabemos que ele pode fazer isso por truques comuns. Por que deveríamos supor que a magia real foi usada?

Essa posição do autor fica mais clara em seu esquadrinhamento sobre o papel dos experimentos de pensamento na produção científica. Segundo ele “qualquer experiência de pensamento pode ser substituída por um argumento sem o caráter de um experimento de pensamento” (Norton, 1996, p. 335 – nossa tradução).

De acordo com Dennett (2012), os experimentos de pensamento são ferramentas importantes para uma ontologia da mente. Há casos em que a linguagem estritamente científica<sup>48</sup>, segundo o autor, dificulta o entendimento e por vezes obscurece os resultados, ou seja, os possíveis equívocos ficam subjacentes ou indetectáveis a princípio. Além disso, no que tange aos tipos de pesquisadores, especialistas e neófitos, tanto cientistas quanto filósofos, se identificam por três tipos, a saber: os gênios, os perseverantes e aqueles que recorrem a recursos diversos para facilitar as suas pesquisas. A esse respeito, Dennett (2012, p. 09 – nossa tradução) comenta:

Algumas pessoas, como von Neumann, são tão gênios naturais que conseguem superar os mais difíceis emaranhados; outros são mais áduos, mas são abençoados com um suprimento heroico de "força de vontade" que os ajudam a manter o curso em sua busca obstinada pela verdade. Depois, há o resto de nós, não prodígios calculistas e um pouco preguiçosos, mas ainda aspirando a entender tudo o que nos confronta. O que podemos fazer? Podemos usar ferramentas de pensamento, às dezenas. Esses práticos extensores protéticos de imaginação e suportes de foco permite-nos pensar de forma confiável e até graciosa sobre questões realmente difíceis.

Para Dennett (2012), filósofos e cientistas recorrem a algumas ferramentas<sup>49</sup> de pensamento mais simples, menos precisas, que abrem mão da necessidade das máquinas sistemáticas da matemática e da ciência empírica, ferramentas manuais da mente dentre as quais podemos destacar: os *rótulos* – essa ferramenta permite criarmos um nome específico, um termo vívido que nos ajuda a mantê-lo sob controle enquanto vamos o examinando em nossa mente (dentre os mais importantes, segundo o autor, estão os de advertência e alarme que nos permitem perceber prováveis fontes de erros); os *exemplos* – apesar da maioria dos cientistas e filósofos o rejeitarem; as *analogias* ou *metáforas* – são ferramentas poderosas, que usamos para delinear as características de algo complexo nos parâmetros de outra coisa complexa que já se entende (ou acredita entender) por completo (mas devemos nos policiar para não desviarmos diante de uma analogia infiel); a *encenação* (staging) – elas possibilitam a formulação de estratégias para resolver problemas articulados e funcionam como uma espécie de prévia ideação; e, finalmente, aqueles extensores protéticos de imaginação que Dennett denomina de bombas de intuição<sup>50</sup>.

---

<sup>48</sup> Segundo Dennett (2012), uma linguagem calcada, quase exclusivamente, na matemática.

<sup>49</sup> Em *Intuition pumps and other tools for thinking*, publicado em 2012, Dennett oferece exemplos dessas “ferramentas”, que nos auxiliam a pensar e aprender sobre ideias complexas. Assim, por exemplo, analogias e metáforas nos permitem a entender e lembrar conceitos complexos ou a perceber e evitar tentações e erros.

<sup>50</sup> Termo cunhado em uma das primeiras críticas ao já citado experimento mental da *sala chinesa* do filósofo John Searle. Inicialmente, alguns críticos entenderam que Dennett os tratou jocosamente, dando-lhe um tom depreciativo, mas o filósofo esclarece a relevância desse tipo de experimento para a filosofia. Todavia, esclarece também que apesar de excelentes ferramentas alguns são de cunho duvidoso e apenas uns poucos são, em sua totalidade, enganosos (cf. Dennett, 2012, p. 11).

Segundo Dennett (1991a; 2012), os experimentos de pensamento estão entre as ferramentas prediletas dos filósofos e cientistas. Ao produzir uma resposta para um problema a partir de uma dedução engenhosa, dispensamos a necessidade de um laboratório. Grandes cientistas recorreram a esse tipo de experimentos e colheram bons resultados – esse fato afasta a tese de que apenas filósofos utilizam tais ferramentas. Com efeito, há EPs que “são analisáveis como argumentos rigorosos, muitas vezes da forma *reductio ad absurdum*, na qual se toma as premissas de seus oponentes e obtém uma contradição formal (um resultado absurdo), mostrando que nem todos podem estar certos” (Dennett, 2012, p. 11 – nossa tradução).

Por conseguinte, as bombas de intuições são experimentos de pensamento que nos permitem testar ideias. São caracterizados por apresentarem menos rigor, mas, segundo o autor, na maioria das vezes, têm grande eficácia. Quase sempre são “pequenas histórias destinadas a provocar uma intuição sincera e impactante – ‘Sim, claro, tem que ser assim!’ – sobre o que quer que seja a tese estará sendo defendida.” (*Ibidem*). Essa visão positiva dos experimentos de pensamento não é compartilhada por uma parte da comunidade científica:

os cientistas muitas vezes suspeitam de discussões teóricas conduzidas em “meras palavras” é que eles reconhecem que a tarefa de criticar um argumento não formulado em equações matemáticas é muito mais complicado e normalmente menos conclusivo. A linguagem da matemática é um confiável executor da contingência. É como a rede na cesta de basquete: remove as fontes de desacordo e julgamento sobre se a bola entrou. (Qualquer pessoa que tenha jogado basquete em uma quadra de playground com um aro nu sabe como pode ser difícil distinguir uma bola de ar de uma cesta.) Mas às vezes os problemas são muito escorregadios e confusos para serem domesticados pela matemática (Dennett, 2012, p. 14 – nossa tradução).

Todo aquele que passou pela faculdade de filosofia, necessariamente se deparou, em algum momento, com experimentos de pensamento clássicos projetados para “bombear” algumas intuições: a parábola da caverna de Platão, em *A República*; o Gênio maligno de Descartes que estabelece a dúvida hiperbólica; o estado de natureza de Hobbes, no qual a vida é desagradável, brutal e curta; são exemplos dentre tantos outros. Ao citar esses exemplos de bombas de intuições Dennett (2012, p. 11 – nossa tradução) reflete sobre os objetivos de seus criadores:

A caverna de Platão pretende nos esclarecer sobre a natureza da percepção e realidade, [...] o demônio maligno é o gerador de ceticismo final, e nosso aprimoramento sobre o estado de natureza quando nós nos contratamos para formar uma sociedade é o ponto da parábola de Hobbes. Estas são as melodias duradouras de filosofia, com o poder de permanência que garante que os alunos se lembrem deles, de forma bastante vívida e com precisão, anos depois de terem esquecido os intrincados argumentos e análises que os cercam.

Nessa perspectiva, consoante com Dennett (2012), devemos considerar os experimentos de pensamento como um recurso com muitas configurações, e devemos explorar todas para ver

se as mesmas intuições ainda são estimuladas quando se considera as variações. Além disso, após a introdução de uma gama de ferramentas gerais, ele canaliza sua análise para o entendimento de como a linguagem e os estados mentais podem adquirir seus significados, a natureza da evolução e da consciência humana.

Considere as seguintes expressões: “água”, “chave” e “mar”. Se, porventura, esses termos fazem algum sentido para o leitor, eles não podem ser tomados apenas como palavras, mas como ferramentas cognitivas abstratas, na mesma proporção em que uma divisão que encontra uma média o é. Cada uma delas têm uma função em um amplo espectro de contextos, facilitando a formulação de hipóteses para serem testadas, simplificando o reconhecimento de padrões imperceptíveis, auxiliando-nos a perscrutar e detectar semelhanças relevantes, e assim por diante (cf. Dennett, 2012, p. 13).

Posto que cada uma das palavras que formam nosso vocabulário seja uma ferramenta, no nível simples, algumas são mais úteis do que outras. Se, por ventura, alguns destes termos úteis não fizerem parte do seu quadro semântico, você tenderá a adquiri-los; em posse de tal ferramenta você poderá articular pensamentos que em circunstâncias contrárias não atingiria o mesmo nível de processamento mental (cf. Dennett, 1991a, p. 234).

Apesar de Dennett (2012) realizar uma explícita apologia aos experimentos de pensamento, é possível observar uma advertência sobre os perigos que uma má aplicação dessa ferramenta pode acarretar, proporcionadas por aquilo que o filósofo denomina de *spear crutches*<sup>51</sup>. No entanto, o autor entende que o rigor matemático adotado por uma parte da comunidade científica – e também por alguns filósofos – não deve ser elevado à enésima potência, haja vista que “há uma hora e um lugar na filosofia para argumentos rigorosos, com todas as premissas numeradas e as regras de inferências mencionadas, mas essas não precisam ser exibidas em público com frequência” (Dennett, 2012, p. 13 – nossa tradução). Dennett (1991a, p. 440 – nossa tradução) trata a questão ainda mais a fundo:

minhas bombas de intuição, como as de qualquer outra pessoa, não são as demonstrações diretas que podem parecer; eles são mais arte do que ciência. [...] Se elas nos ajudarem a conceber novas possibilidades, que podemos então confirmar por métodos mais sistemáticos, isso é uma conquista; se elas nos atraem pelo caminho das flores, é uma pena. Mesmo boas ferramentas podem ser mal utilizadas e, como

---

<sup>51</sup> Muletas de lança é a denominação dennettiana para os casos em que a aplicação das bombas de intuição foram mal empregadas, ou seja, “o termo [...] para ferramentas de pensamento que saem pela culatra, aquelas que só parecem ajudar na compreensão, mas que, na verdade, espalhou escuridão e confusão em vez de luz.” (Dennett, 2012, p. 15). De acordo com o autor, o experimento mental do quarto chinês, desdobrado por Searle, é um caso típico dessa má aplicação, pois, essa noção ignora algumas faculdades inerentes à consciência – memória, conhecimento de experiências de fato, racionalidade, etc. –, bem como é insuficiente para descrever todos os níveis de IA (cf. Dennett, 1991a, p. 438).

qualquer outro trabalhador, faremos melhor uso se entendermos como nossas ferramentas funcionam.

Tendo em vista suas opiniões sobre o tema, fica claro que para Dennett (2012) os experimentos de pensamento ocupam um papel relevante na produção de conhecimento. No entanto, a necessidade de estabelecer uma taxonomia para esse tipo de recurso metodológico não proporciona uma maior eficácia na sua aplicação. Quando Brown executa uma classificação para os EPs, ele tem a finalidade de estabelecer uma defesa de sua epistemologia – calcada no conceito de experimento de pensamento platônico. Não obstante, esse objetivo não pode ser rotulado como fútil, como faz Norton, pois ele nos ajuda a entender e a aplicar esse tipo de ferramenta. Dennett vai mais além. O filósofo está convencido de que não se pode fazer ciência – e muito menos filosofia – sem recorrer às bombas de intuição. Segundo o autor, são a elas que recorreremos quando a linguagem matemática, preferida pela comunidade científica, é insuficiente para abordar determinado problema.

A aplicação de experimentos de pensamento no campo da filosofia da mente é muito frequente e podem ser usadas tanto para refutar quanto corroborar uma teoria. No capítulo III, iremos examinar o grau de eficácia dessa ferramenta, analisando dois exemplos famosos de utilização de EPs no campo da filosofia da mente (delineados por John Searle e David Chalmers) para ilustrar como eles podem nos ajudar a produzir uma ontologia da mente, bem como podem nos levar a conclusões equivocadas.

## 2. UMA TEORIA EMPÍRICA DA CONSCIÊNCIA

No capítulo I, lançamos luz a algumas disciplinas que subvenciam o pensamento dennettiano. Elas são imprescindíveis para assimilarmos os conceitos basilares que permeiam a sua tese. Neste capítulo, iremos nos dedicar ao exame desses conceitos, a fim de que possamos realizar um esquadrinhamento nos parâmetros exigidos para nos apropriarmos amplamente do funcionalismo dennettiano. Essa aproximação das ciências cognitivas leva, como veremos, a uma nítida divergência em relação ao behaviorismo skinneriano, pois culmina na ideia de superação da psicologia do comportamento de Skinner.

Portanto, segundo Dennett (1991a; 2006), devemos não só recorrer aos trabalhos desenvolvidos na ciência cognitiva, mas também às interfaces que abrangem justamente as pesquisas que proporcionam um caráter epistemológico para sua teoria. Dito de outro modo, a teoria dennettiana consiste em desenvolver uma filosofia da mente que tenha como formação de base os trabalhos desenvolvidos no campo das ciências empíricas, intenciona solucionar a questão da ontologia da consciência sem a necessidade de uma teoria da consciência e de uma teoria do conteúdo. Todavia, de acordo com o filósofo, se quisermos edificar uma teoria da mente genuinamente empírica, faz-se necessário sobrepujar o dualismo cartesiano. Com efeito, delinear uma crítica contundente ao dualismo da substância cartesiana culminará em uma revisão da abordagem fisicalista da mente, pois a adoção do fisicalismo, em detrimento do dualismo da substância, não é suficiente para eliminar os problemas da relação mente/cérebro.

Nesse sentido, é preciso sair da zona de conforto que nos proporciona a psicologia popular e abandonar a ideia de que existe um local específico no cérebro onde se manifesta a consciência. Uma vez estabelecida a impossibilidade descritiva, na abordagem dualista, da interação entre a mente imaterial e o cérebro “talvez algum alargamento básico da ontologia das ciências físicas seja necessário para explicar os fenômenos da consciência” (Dennett, 1991a, p. 36 – nossa tradução).

Em detrimento ao teatro cartesiano, qual seria a alternativa? Se não posso considerar a extravagante ideia de que a mente é uma substância não física – que interage com um cérebro físico – e não devo admitir que cada estado mental é um tipo de estado cerebral correspondente, qual seria então a terceira via?

Considere um sistema simples específico – uma espécie de vida unicelular, por exemplo. Dado seu caráter de espécie viva, esse organismo se configura por uma série de ações fenotípicas (comportamentos limitados a sua espécie) que visam a manutenção da sua existência e de seus descendentes. O fato pelo qual, em uma situação que lhe é oferecido simultaneamente

um tipo de ácido corrosivo e um tipo de proteína, sua tendência em evitar o primeiro e se deslocar em direção ao segundo nos parece estar em consonância com aquilo que esperamos de uma espécie viva e, agora, não nos absteríamos de lhe atribuir certo grau de inteligência ou intencionalidade<sup>52</sup>.

Podemos agora, na esteira de Dennett (1991a), realizar uma analogia entre sistemas e nossas mentes. Temos a tendência de dissociar as semelhanças de nossos comportamentos atrelados aos estados mentais e ao comportamento, por exemplo, de organismos unicelulares, mas quando examinamos os detalhes de seu funcionamento mais de perto, percebemos um alto grau de semelhança. De acordo com o estadunidense, nossos cérebros modulam informações e traçam objetivos. Essas informações, que configuram a mente, são fornecidas por neurônios e conjuntos de neurônios<sup>53</sup>. Dito de outro modo, é como se esses neurônios “fossem agentes diminutos dotados de uma mente simples, serviços especializados racionalmente procurando atingir seus objetivos obsessivos particulares ao agir de maneiras ditadas por sua percepção das circunstâncias” (Dennett, 1997, p. 24). O tipo de entidade que se configura nesses termos, sendo ela uma elevação que vai das mais simples às mais complexas, é rotulada de sistemas intencionais e o comportamento oriundo de tais sistemas (pseudo ou genuínos) de postura intencional.

As regras básicas da teoria dennettiana são diretas: a princípio, deve-se desconsiderar o tecido *Wonder* (o aspecto maravilhoso). O caráter intrigante da consciência será explicado dentro dos parâmetros da ciência física contemporânea, ou seja, será evitado abordar o problema recorrendo a forças, substâncias ou poderes orgânicos miraculosos, sem explicação empírica. Além disso, objetiva verificar o que pode ser realizado dentro dos limites conservadores de nossa ciência padrão, evitando a solução do problema mente/cérebro com aspectos promissórios de uma ciência futura (cf. Dennett, 1991a, p. 40).

Na sequência, Dennett (1991a) procura escamotear-se do rótulo de “anestesiado”. Tal como é cediço, uma das principais críticas dirigidas aos behavioristas<sup>54</sup> é a negação do testemunho das experiências introspectivas – fingem estar anestesiados –, desconsiderando, portanto, como método propedêutico, a observação dos estados mentais em primeira pessoa.

---

<sup>52</sup> O termo aqui empregado não deve ser confundido com a concepção husserliana – descrita por Dennett como fenomenologia pura. Neste capítulo dedicamos uma seção para tratar desse conceito segundo a perspectiva dennettiana.

<sup>53</sup> Segundo Dennett (1991a), há vários exemplos desse tipo de entidades – que variam do nível molecular ao continental. Além disso, não se resumem a seres naturais, como animais, vegetais e seus respectivos sistemas, mas a entidades artificiais. O despertador é um exemplo de um “pseudo agente simples”.

<sup>54</sup> Em Dennett (1991a; 2006), o conceito de heterofenomenologia apresenta uma nítida dependência do behaviorismo clássico.

Posto que o behaviorismo tenha surgido em resposta ao dualismo, não pode ser uma alternativa plausível, pois executa uma desqualificação radical do mentalismo, ou seja, afirmam que os atributos mentais – desejos, crenças, etc. – não se baseiam em dados empíricos, não são passíveis a experimentos e, portanto, não podem ser cientificamente comprovados (cf. Dennett, 2006, p. 97). Os psicólogos behavioristas entendem que “os estados e eventos internos, sejam neurofisiológicos ou mentais, não são objetos apropriados de investigação psicológica – os estados e eventos neurofisiológicos são assuntos dos biólogos, os estados e eventos mentais, se é que eles existem, não devem ser seriamente mencionados” (Bunnin; Tsui-James, 2004, p. 167). Destarte, os estados mentais são prescindidos e eliminados por serem entendidos como excrescências metafísicas.

Nesse sentido, de acordo com Dennett (1991a) só é possível realizar as rejeições das características controversas da consciência se for possível demonstrar que de algum modo são ilusórias, algo análogo a um truque de mágica. Além disso, evita calcar seu argumento em descobertas recentes da ciência, de sorte que procura “obter todos os fatos científicos corretos, na medida em que são conhecidos hoje, mas há muita controvérsia sobre quais avanços empolgantes resistirão ao teste do tempo” (Dennett, 1991a, p. 40 – nossa tradução). Por conseguinte, não se trata aqui de uma negação sumária da totalidade de nossa experiência consciente, mas evitar a nossa tendência de “pensar que somos muito mais imunes a erros do que somos” (Dennett, 1991a, p. 68 – nossa tradução). Essa é a postura do funcionalismo, em detrimento à psicologia popular – que tem como grande limitação a descrição do testemunho da experiência em primeira pessoa sem examinar, ou sequer considerar, os fundamentos que proporcionam as causas e os efeitos da experiência consciente – herdeira direta do dualismo da substância. Devemos nos policiar quanto ao excesso de confiança dado ao testemunho de nossos estados mentais particulares.

Portanto, de acordo com Dennett (2006), uma teoria da mente não pode estar em conflito com as teorias neurofisiológicas, nem psicológicas, ou ter a intenção de substituí-las, entretantes deve fundamentá-las e complementá-las – deve desenvolver justificativas para estabelecer os pressupostos metafísicos fundamentais – e, ao que parece, todo neurocientista deveria estar cômico disto, pois:

Todo neurocientista sabe que mesmo na idade de ouro do conhecimento neurofisiológico, quando a atividade de todo seguimento de fibra estiver bem compreendida, questões como “*o que é a consciência?*” e “*o que há nas dores que as faz repulsiva?*” não terão respostas em seus livros-texto – a não ser que eles incluam capítulos de filosofia (Dennett, 2006, p. 15 – grifo nosso).

Antes de apresentarmos os conceitos basilares para as reflexões de Dennett acerca da natureza da inteligência consciente, retomemos a análise, um pouco mais aprofundada, da corrente dualista proposta por René Descartes – que foi uma das críticas realizadas pelo autor para fundamentar sua teoria –, a chamada teoria da substância, bem como expor a crítica direcionada ao fisicalismo “ingênuo”.

## 2.1 Crítica ao dualismo da substância e ao fisicalismo<sup>55</sup> ingênuo

Faria mais sentido afirmar que “eu possuo um corpo” ou que “eu sou meu corpo”? Para a grande maioria das pessoas parece óbvio que possuímos um corpo. Esse entendimento é compreensível, já que nós, enquanto “alma”, nos relacionamos com o mundo através dele. Nele, e por ele, percebemos não só a densidade das coisas, seus aspectos tridimensionais – altura, largura e comprimento – e a sua degradação, mas também a relação do nosso corpo com nossa mente. Assim, para a psicologia popular, o corpo é sempre referenciado como algo distinto da mente. No mais das vezes, se alguém machuca a mão certamente dirá “machuquei a minha mão”, ou seja, essa parte do corpo é tomada como uma posse. Essa expressão deixa tácitos dois entendimentos: o primeiro emerge da ideia de que a identidade do indivíduo, a sua *self*, não está relacionada diretamente com o corpo, o seu “Eu” seria a sua mente; o segundo, apesar de ser uma derivação do primeiro entendimento, traz a noção de que a mente é algo extremamente distinto do corpo. Além disso, essa concepção da relação mente-corpo dá margem ao entendimento que, diferentemente do corpo, a mente não possui extensão e, por assim dizer, somos um fantasma na máquina<sup>56</sup> (cf. Dennett, 1991a, p. 29).

As pesquisas acerca do funcionamento do cérebro auxiliam as abordagens em relação ao problema ontológico da mente. Nesse sentido, o problema alma<sup>57</sup>/corpo, levantado por Descartes, se converte no problema mente/cérebro. Entrementes, a essência do problema não muda: aquele que se dispor a refutar o dualismo da substância deve explicar como o pensamento emana do cérebro e como é possível esse mesmo pensamento influenciar o próprio cérebro.

---

<sup>55</sup> O fisicalismo pode ser descrito como a tese metafísica que resume tudo ao físico. Reduz tudo a uma superveniência lógica, de modo que apresenta um monismo ontológico de substância, se opondo assim ao dualismo. Todavia, grande parte dos fisicalistas admitem que a natureza pode conter fenômenos, que a princípio, não parecem participar do mundo físico — como por exemplo, a mente e a consciência. Enquanto os dualistas, como David Chalmers por exemplo, entendem que a mente/consciência e o cérebro são duas coisas categoricamente distintas, os fisicalistas (Searle e Dennett) reiteram que, em última análise, os estados mentais são de cunho físicos ou são supervenientes a uma estrutura física específica. Invariavelmente o fisicalismo é confundido ou associado à corrente materialista, por apresentar, em pesquisas relacionadas ao estudo da mente, uma noção mais ampla de materialismo.

<sup>56</sup> A expressão “o dogma do fantasma na máquina” foi criada por Gilbert Ryle (1900-1976) para deflagrar uma crítica ao conceito de *res cogitans* de Descartes (cf. Dennett, 1991a, p. 33).

<sup>57</sup> Termo utilizado por Descartes para se referir a mente (cf. Descartes, 2001, p. 39).

Diante da perspectiva que distingue o cérebro e a mente emergem algumas questões: se mente e cérebro são coisas diferentes, qual seria o fio que estabelece essa ligação? Por que danos cerebrais influenciam rigorosamente no funcionamento da mente? O dualismo apresenta extrema dificuldade para responder essas questões.

Em consonância com Tripicchio (2004), uma das inerências da filosofia moderna foi a reviravolta condicionada pela reflexão filosófica em direção ao *sujeito cognoscente*. Ao estabelecer a *dúvida hiperbólica* como método para alcançar sua primeira certeza – o *cogito* cartesiano –, René Descartes torna-se um dos grandes nomes do racionalismo. Seu objetivo é chegar à verdade por meio da demonstração e do exame – mas essa atenção dada ao sujeito cognoscente proporcionou um novo jeito de problematizar a conexão entre a alma e o corpo. Assim, já no século XVII, os pensadores perscrutam o entendimento de como pode haver uma relação causal entre algo não-físico com uma substância rigorosamente material. O novo que surge com o filósofo francês foi a iniciativa de uma investigação acerca de como a alma pode interagir causalmente com o corpo.

Os racionalistas a partir do conceito de substância, procuram problematizar a relação alma/corpo, abandonando os dogmas do pensamento religioso. Vejamos, por exemplo, como Descartes (2001, p. 39) aborda essa questão:

Depois, examinando atentamente o que eu era e vendo que podia fingir que não tinha nenhum corpo e que não havia nenhum mundo, nem lugar algum onde eu existisse, mas que nem por isso podia fingir que não existia; e que, pelo contrário, pelo próprio fato de eu pensar em duvidar da verdade das coisas, decorria muito evidentemente e muito certamente que eu existia; ao passo que, se apenas eu parasse de pensar, ainda que tudo o mais que imaginara fosse verdadeiro, não teria razão alguma de acreditar que eu existisse; por isso reconheci que eu era uma substância, cuja única essência ou natureza é pensar, e que, para existir, não necessita de nenhum lugar nem depende de coisa alguma material. De sorte que este eu, isto é, a alma, pela qual sou o que sou, é inteiramente distinta do corpo, e até mais fácil de conhecer que ele, e, mesmo se o corpo não existisse, ela não deixaria de ser tudo o que é.

Como vimos, Descartes (2001) concebe uma ontologia que estabelece o homem como um ser caracterizado por uma dualidade de substâncias: a coisa pensante e a coisa extensa. De acordo com o filósofo francês, essas substâncias influenciam-se mutuamente. A teoria da substância (como todo dualismo derivado dela) faz uso de um forte apelo intuitivo, já que de modo introspectivo, nossa mente não se assimila como algo físico. No entanto, a grande dificuldade do cartesianismo, característica de toda corrente dualista, é de responder como algo imaterial que é a mente, na sua suposta supremacia em relação ao corpo, pode produzir efeito causal sobre ele. A esse respeito Churchland (2004, p. 28) nos dá mais detalhes:

Como o próprio Descartes percebia (ele foi um dos primeiros a formular a lei da conservação da quantidade de movimento) a matéria no espaço se comporta de acordo com leis rígidas, e não se pode obter movimento corporal (= quantidade de

movimento) a partir do nada. Como essa substância pensante totalmente incorpórea pode ter influência sobre a matéria dotada de peso? Como podem duas coisas tão diferentes ter algum tipo de contato causal? Descartes sugeria que uma substância material sutil – os espíritos animais – transmitia a influência da mente ao corpo em geral.

A resposta de Descartes (2001) é no mínimo intrigante. Ele alvitra a existência de uma substância material<sup>58</sup> bem mais sutil, um elemento que realizaria esse papel de mediador que estabeleceria a influência da alma ao corpo (cf. Churchland, 2004, p. 28). Todavia, não é preciso nos aprofundarmos na questão para perceber que esse argumento não resolve o problema. Ainda que fosse possível admitir essa substância, por mais sutil que ela seja, de alguma maneira (ou momento) seria material, portanto, a pergunta permanece não respondida. Desse modo, a reflexão cartesiana foi insuficiente para identificar, com precisão, a suposta linha tênue entre mente e corpo. A respeito do princípio da conservação de energia, de forma bastante eloquente, Dennett (1991a, p. 35 – nossa tradução) contribui conosco:

Um princípio fundamental da física é que qualquer mudança na trajetória de qualquer entidade física é uma aceleração que requer gasto de energia, e de onde virá essa energia. É este princípio de conservação de energia que explica a impossibilidade física de “máquinas de movimento perpétuo”, e o mesmo princípio é aparentemente violado pelo dualismo. Este confronto entre a física bastante padrão e o dualismo tem sido continuamente discutido desde os dias de Descartes, e é amplamente considerado como a falha inevitável e fatal do dualismo.

Descartes estava cômico deste fato – pois, foi um dos que contribuiu para a ideia de movimento. Ele sabia que o comportamento da matéria no espaço obedece a leis rigorosas, sendo que uma delas é que não se pode obter movimento perpétuo, ou seja, movimentar-se a partir do nada. Não precisamos ir muito longe para perceber que sua tese não logrou êxito. Fazendo emergir um problema incorrigível na teoria da substância: como pode a substância pensante proporcionar uma relação causal com um corpo material?

O dualismo da substância dá margem a novas análises e correções, possibilitando o surgimento de outras linhas de abordagem, tais como, o dualismo popular e o dualismo da propriedade<sup>59</sup> – que possui pelo menos duas outras perspectivas: o epifenomenalismo e o dualismo interacionista – para citar alguns.

Contudo, como a base, seus pressupostos, para as abordagens das ramificações dualistas são os chamados legados cartesianos: 1) que a alma é conhecida intuitivamente; 2) que alma se ocupa da manipulação do corpo; 3) e que há um problema teórico de entender como nós, seres

---

<sup>58</sup> Descartes estava equivocado ao acreditar que a glândula pineal poderia ser uma das conexões entre o cérebro e a consciência. Não há como precisar o momento exato onde cada evento consciente se dá, pois, aparentemente a cognição e o controle são distribuídos por todo o cérebro (cf. Dennett, 1991a, p. 169).

<sup>59</sup> Segundo essa perspectiva, o cérebro é dotado de propriedades especiais de caráter não-física (cf. Churchland, 2004, p. 31).

conscientes, podemos garantir que o mundo objetivo existe de fato, ainda que tenhamos soluções sustentáveis para esse problema – e como “até recentemente, de fato, a filosofia da mente foi dominada pela perspectiva cartesiana da ‘primeira pessoa’ ou por uma perspectiva de dentro, interna” (Bunnin; Tsui-James, 2002, p. 165), o pensamento dennettiano concentra seu esforço, num primeiro momento, na refutação desse tipo de concepção e em seguida estender essa crítica a toda corrente dualista. Qual seria então o problema da teoria cartesiana? E por que houve (deve haver) uma ruptura com a perspectiva dualista?

A doutrina dualista se mostra, aparentemente, pouco fecunda, pois, não se ocupa em descrever a natureza da mente, mas tão-somente em afirmar que ela se distingue do corpo, ou seja, não nos diz o que a mente é, mas se limita a dizer o que ela não é, procura simplesmente estabelecer a sua existência a partir de uma “desanalogia” com a matéria. Dito de outro modo, a abordagem dualista assume um caráter negativista, ou seja, procura descrevê-la a partir da extração das características das coisas materiais.

Nessa perspectiva, consoante com Dennett (1991a), o apelo ao testemunho da introspecção nos leva, quase que inevitavelmente, à adoção do dualismo. Mas apesar de serem muitos os argumentos utilizados para defender o dualismo, os problemas se proliferam. Se quisermos nos afastar dessa concepção de teoria da mente, devemos estar dispostos a conceber uma nova perspectiva de análise da coisa-mente.

No século XX, houve uma rápida mudança de perspectiva de análise da relação mente-cérebro por dois motivos: o primeiro foi o grande impacto promovido pelo Círculo de Viena, o denominado positivismo lógico, e seu critério de verificabilidade. O significado linguístico e a probidade científica passaram a ser formulados mediante a testabilidade intersubjetiva. Para que a mente fosse tomada como objeto de estudo científico, suas atribuições teriam que permitir testes publicamente verificáveis. Os objetos da ciência, de modo geral, assumem uma perspectiva intersubjetiva, uma análise em “terceira pessoa”. Como consequência disso, “a tradição da perspectiva da ‘primeira pessoa’ teve de ser abandonada para fins científicos e, sentia-se, para fins metafísicos” (Bunnin; Tsui-James, 2002, p. 166). O horizonte das pesquisas seria agora essa “avenida que foi explorada com entusiasmo e frutuosidade durante o último quarto de século sob as bandeiras gêmeas da ciência cognitiva e do funcionalismo – a extrapolação do naturalismo mecanicista do corpo para a mente” (Dennett, 1999, p. 05 – nossa tradução).

Por conseguinte, o surgimento de uma gama de objeções estruturais inerentes ao dualismo leva as pesquisas a um novo patamar. Não se podia mais fazer vistas grossas ao desacordo da abordagem incorpórea da mente com a concepção física e científica – embora o

conhecimento da mente não seja fundamental para os trabalhos da neurofisiologia, que explica os movimentos dos membros corporais – e, partindo do pressuposto que a evolução se deu, exclusivamente, mediante processos de seleção natural e físicos de mutação de organismos unicelulares (sem vestígios de algo que constitua uma mente) até a espécie humana “é anômalo supor que, em algum momento, a *Mãe Natureza* (na forma da genética das populações) tenha criado, de algum modo, as mentes incorpóreas cartesianas, além das células e dos órgãos físicos” (Bunnin; Tsui-James, 2002, p. 166 – grifo nosso). Além disso, se de fato houvesse um ente que possuísse as prerrogativas delineadas pelos dualistas – emoções, raciocínios e consciência – mas que, ao mesmo tempo, dependesse do cérebro para ter as experiências sensoriais (entrada de dados) que fundamentam as suas volições e comportamentos, então “a razão, a emoção e a consciência seriam invulneráveis ao controle direto ou às patologias resultantes da manipulação ou dano ao cérebro” (Churchland, 2004, p. 45). O problema é que não se observa nada semelhante a isto. Estava aberto o espaço para uma nova abordagem: o fisicalismo.

Não obstante, Dennett (1991a) propõe um exame dos argumentos de algumas abordagens de cunho fisicalista<sup>60</sup>, pois esse tipo de análise também apresenta problemas. Essa corrente tem a sua maior representação na *teoria da identidade* (materialismo reducionista) que parte do pressuposto de que existe uma correspondência entre atividades cerebrais e estados mentais. O grande problema da teoria da identidade é que os estados mentais (desejo, dor, etc.) não parecem ser radicalmente análogos aos estados neurofisiológicos, ou seja, há um problema nas concepções que concebem um neurônio para cada estado mental – e esse problema é exclusivamente filosófico. Acerca do fisicalismo ingênuo, Dennett (1991a, p. 106 – grifo nosso) esclarece que:

embora o materialismo de um tipo ou de outro seja agora uma opinião aceita que se aproxima da unanimidade, mesmo os materialistas mais sofisticados hoje muitas vezes esquecem que, uma vez que a *res cogitans* fantasmagórica de Descartes é descartada, não há mais um papel para um portal centralizado, ou mesmo para qualquer centro funcional para o cérebro.

Apesar de que essa limitação do fisicalismo seja uma evidência – conceber a mente, os estados mentais, a partir da perspectiva dualista, ou seja, baseado na introspecção –, ainda é considerado por muitos a melhor maneira de iniciar uma pesquisa sobre a inteligência consciente, pois, diferentemente da abordagem dualista, está calcado nos trabalhos da neurociência, da psicologia cognitiva, da etologia e da ciência da computação. Essa

---

<sup>60</sup> Essa crítica engloba as primeiras teorias materialistas.

aproximação das ciências empíricas proporciona uma vantagem aos fisicalistas na edificação de teorias.

O estado consciente é o efeito da atividade cerebral – o cérebro é sua causa, dirá o fisicalista. Não obstante, examinar o problema da mente a partir da análise em primeira pessoa não pode render bons frutos, ou seja, do ponto de vista introspectivo não podemos examinar os processos subjacentes à estrutura dos estados mentais sem recair em uma abordagem dualista inconsciente e sucumbir aos problemas a ela inerentes. Emerge aqui a necessidade de examinar um dos conceitos delineados por Dennett para resolver esse problema, a saber: a heterofenomenologia.

## 2.2 O método heterofenomenológico

De acordo com Dennett (1991a), o teatro cartesiano é extremamente sedutor por supostamente preservar a diferença entre realidade e aparência no âmago da subjetividade humana. Entretanto, é pouco motivador cientificamente por apresentar um alto grau de especulação, porquanto apresenta a estranha categoria do objetivamente subjetivo, ou seja, os modos pelos quais “as coisas, na verdade, objetivamente parecem para você, mesmo que elas não pareçam assim para você” (Dennett, 1991a, p. 132 – nossa tradução). Somente um método que se apresente como guardião da abordagem científica – que tenha como princípio a observância da busca dos resultados das ciências empíricas – poderia garantir uma descrição dos estados mentais bem mais aceitável. Por ora, vejamos o que Dennett (1991a, p. 72 – nossa tradução) diz a esse respeito:

um caminho neutro que leva da ciência física objetiva e sua insistência no ponto de vista da terceira pessoa, a um método de descrição fenomenológico que pode (em princípio) fazer justiça ao subjetivo mais privado e inefáveis experiências, sem nunca abandonar os escrúpulos metodológicos da ciência.

Dennett (1991a) herda do behaviorismo o método de análise em terceira pessoa, ou seja, utiliza o behaviorismo<sup>61</sup> filosófico como método. Assim, evita aderir às implicações ideológicas dessa abordagem – caracterizada por alguns princípios – cuja neurociência também experimenta como pressuposto. Esses princípios podem ser elencados da seguinte maneira: 1) estados mentais são inexistentes<sup>62</sup>; 2) estados mentais existem, no entanto, não exercem

---

<sup>61</sup> O behaviorismo *filosófico* aqui discutido deve ser nitidamente distinguido do behaviorismo *metodológico* que teve uma influência muito grande no âmbito da psicologia. Em sua forma menos sofisticada, “o behaviorismo metodológico insiste em que todo novo termo teórico inventado pela ciência da psicologia *deve ser* definido em termos operacionais, a fim de garantir que a psicologia mantenha contato firme com a realidade empírica.” (Churchland, 2004, p. 52).

<sup>62</sup> Essa posição é comumente rotulada de *Behaviorismo Descalço*.

nenhuma influência no comportamento<sup>63</sup> – não podem ser estudados pelas ciências empíricas; 3) estados mentais existem e influenciam no comportamento<sup>64</sup>. Não obstante, essa influência não abre precedente para um exame científico, restando, portanto, a satisfação com as hipóteses sobre esses “efeitos e processos periféricos ou inferiores no cérebro” (Dennett, 1991a, p. 71 – nossa tradução).

Decerto, cada uma dessas concepções está em consonância com o pensamento de Thomas Nagel segundo o qual há uma defesa gradual de um solipsismo incorrigível, a partir de uma crítica do materialismo reducionista (cf. Nagel, 2005, p. 252). Esse ceticismo é colocado à prova na edificação do método heterofenomenológico dennettiano.

A aplicação do método heterofenomenológico, segundo Dennett (1991a), envolve o conceito de *Sujeito Heterofenomenológico*, um ente ficcional a quem atribuímos “de fora” crenças e desejos experimentados. Esse sujeito é apenas uma ficção teórica<sup>65</sup>. Não seria possível existir um sujeito real que experimente essas experiências de fato? Um sujeito segundo o qual o cérebro monta o palco para a exibição de tais experiências? Podemos realmente nos abster do teatro cartesiano? Em relação a essa última indagação, o filósofo está convicto que sim.

Dennett (1991a) lança mão a um experimento de pensamento para esclarecer a abordagem heterofenomenológica – que nos permite apresentar dados contra-intuitivos, a saber: o experimento mental do antropólogo.

Considere um grupo de antropólogos que têm contato com uma tribo até então desconhecida. A tribo em questão acredita em um deus da floresta, denominado Feenoman. Ao tomar conhecimento desse deus, os antropólogos estão diante de um dilema: se converter sinceramente a religião e adotar todos os ritos de fé que envolve essa crença; ou podem assumir uma postura agnóstica para estudar os rituais de culto. Assumindo a postura agnóstica é possível para esses antropólogos, mesmo não crendo em Feenoman, colher dados dessa religião e sistematizá-los da melhor maneira. Eles podem tentar estabelecer uma descrição de Feenoman a partir dos relatos dos nativos. Procuram um ponto de convergência, mas nem sempre é possível, pois, alguns podem descrever Feenoman como um ser de olhos azuis e outros o descrevem com olhos verdes. Alguns, ainda, podem tomá-lo como uma deusa. Os antropólogos

---

<sup>63</sup> O *epifenomenalismo* é uma variante desse tipo de abordagem (cf. Churchland, 2004, p. 31).

<sup>64</sup> Dennett (1991a; 1991b) esclarece que essa posição é amplamente aceita pelos neurocientistas – sobretudo, por aqueles que nutrem asco aos teorizadores da consciência. A rigor, trata-se de um dualismo, pois, esse grupo apresenta uma concordância com Descartes por entender que a mente não é o cérebro, comungam da ideia de que devemos aceitar apenas uma teoria do cérebro. No entanto, esse princípio é o que mais se aproxima da tese dennettiana.

<sup>65</sup> Construções teóricas abstratas (cf. Teixeira, 2008, p. 16).

se esforçam por dirimir essas divergências identificando e rejeitando os relatos de má-fé – além de promover reformulações dos relatos mais sinceros. Essa metodologia dos antropólogos implica em uma gradual emergência de uma descrição mais lógica: a crença de um povo em um deus (Feenoman) que possui um conjunto de características “físicas” e hábitos específicos, que preza por determinados comportamentos de seus fiéis e que rejeita outros. Por outro lado, os crentes em Feenoman o tomam como uma entidade real e não como um mero objeto intencional. O próprio Feenoman os instruiu sobre sua natureza – o que, acredita-se, não deixa margem para duvidar de sua existência (cf. Dennett, 1991a, p. 82). A postura dos crentes é análoga aos que executam a análise em primeira pessoa para investigar a natureza da inteligência consciente. A postura dos antropólogos é heterofenomenológica.

Segundo Dennett (1991a), utilizar os relatos das pessoas para tentar entender o que ocorre em suas mentes, aplicar o método heterofenomenológico, é muito mais auspicioso para uma teoria da mente. Caracteriza-se por estabelecer uma reconstrução de relatos subjetivos de outras pessoas – uma remontagem que, apesar de desconsiderar análise em primeira pessoa de seus próprios estados mentais, permite a realização de interpretações desses relatos numa perspectiva intencional. Como método, a heterofenomenologia rejeita os dados fornecidos pelo acesso exclusivo de nossos próprios estados mentais, privilegiando as descrições de dados mediada pela linguagem (cf. Teixeira, 2008, p. 24).

A esta altura o leitor já percebeu que o método heterofenomenológico é de caráter contra-intuitivo. Ele assume uma postura neutra em relação às alegações experimentais dos sujeitos (não contesta e nem as aceita como verdadeiras), ou seja, assume uma postura de análise construtiva com o intuito de elaborar uma descrição fiel do mundo segundo os relatos dos sujeitos, sem que os sujeitos em questão não percebam essa neutralidade (cf. Dennett, 1991a, p. 83). A esse respeito Dennett (1991a, p. 98 – nossa tradução) nos oferece mais detalhes quando afirma que:

desenvolvemos um método neutro para investigar e descrever a fenomenologia. Envolve extrair e purificar textos de (aparentemente) sujeitos falantes e usar esses textos para gerar uma ficção teórica, o mundo heterofenomenológico do sujeito. Este mundo ficcional é povoado por todas as imagens, eventos, sons, cheiros, palpites, pressentimentos e sentimentos que o sujeito (aparentemente) acredita sinceramente que existam em seu (ou nosso) fluxo de consciência.

Consoante com Dennett (1991a), devemos ser flexíveis em admitir se os relatos dos sujeitos aparentes são oriundos de mentirosos, zumbis ou um alienígena travestido de ser humano – se percebido, não há necessidade de provocá-los temeridade anunciando esses fatos. Vale ressaltar que essa postura de neutralidade é de caráter efêmera, uma fase no

desdobramento de uma teoria efetivamente empírica que deve contemplar os estados mentais de tais sujeitos.

Além disso, segundo Dennett (1991a; 1995a; 1998), para entendermos as fontes de *design* da consciência humana é preciso examinar o modo pelo qual a cultura se transformou em um repositório e meio de transmissão de inovações – não só da consciência, pois, como veremos, ela mesma é mais um meio de evolução.

### **2.3 A tese Memética: a consciência como um *pandemonium***

A conveniência da eliminação do teatro cartesiano enaltece a teoria funcionalista homuncular dennettiana, mas essa rejeição não é um fardo que se possa jogar fora sem a ajuda da neurociência<sup>66</sup>. Emerge aqui um aspecto importante da teoria dennettiana, a saber: a teoria memética – uma teoria que, em conjunto com as pesquisas da neurociência, descreve o funcionamento da mente humana. Essa postura dificulta uma infirmação ao funcionalismo dennettiano, haja vista que é impossível realizar uma análise do cérebro verdadeiramente científica em primeira pessoa. A esse respeito, Dennett (1991a, p. 94 – nossa tradução) trata a questão ainda mais a fundo:

há circunstâncias em que as pessoas estão simplesmente erradas sobre o que estão fazendo e como estão fazendo. Não é que elas estejam na situação experimental, mas que elas confabulam; elas preenchem as lacunas, adivinham, especulam, confundem teorização com observação. A relação entre o que elas dizem e o que as leva a dizer o que dizem dificilmente poderia ser mais obscura, tanto para nós heterofenomenólogos de fora quanto para os próprios sujeitos. Eles não têm como “ver” (com um olho interior, presumivelmente) os processos que governam suas afirmações, mas isso não os impede de ter opiniões sinceras para expressar.

Dennett (1991a) está convicto que a pressuposição de um administrador central em nosso cérebro é infecunda. Comumente admite-se que há um administrador no centro do processo mental encarregado de articular a comunicação com o mundo exterior. Para conceber a ideia de que o outro possui uma mente é preciso considerar que haja um “eu”, um burocrata, um coordenador de todos os sentidos. Quando interpretamos um argumento relativamente coerente, não conjecturamos que sejam murmúrios desconexos ou palavras desarticuladas, mas que há um agente administrativo extremamente consciente. Todavia, segundo o filósofo, não ocorre nada parecido em nossos cérebros. Não existe uma hierarquia administrativa, similar a

---

<sup>66</sup> Se partirmos do pressuposto de que o processo visual, por exemplo, é uma simples retransmissão, pois, se a imagem captada de modo invertida se faz necessário um segundo processo para torná-la compatível com o real, ao longo do nervo ocular para o *lobo occipital* (área visual do cérebro), cometeremos um erro abissal: a chamada falácia do homúnculo – um processo de reinterpretação de imagens realizada por um “homenzinho” dentro de nossas cabeças. Mas, a rigor, esse “homenzinho” teria também o seu homúnculo, e esse último também o seu, de tal modo que o processo se configura em uma regressão ao infinito de processos visuais. Esse entendimento não resolveria o problema da percepção, menos ainda o da consciência (cf. Ramachandran, 2014, p. 74).

isso no cérebro, que chefia a produção de fala ou escrita. O que pôr em seu lugar? Se não há o grande administrador, como se estabelece o significado? Sua substituição se dará por meio da noção de um darwinismo neural<sup>67</sup>, uma luta para oferecer o melhor pensamento/resposta que conduzirá o indivíduo a um comportamento correspondente.

Um dos conceitos basilares – e um dos mais polêmicos – da teoria da mente dennettiana é o conceito de memes<sup>68</sup>. Ele nos permite entender “a maneira pela qual a cultura se tornou um meio de repositório e transmissão para inovações (não apenas inovações de consciência) e é importante para entender as fontes de design da consciência humana, pois, é mais um meio de evolução” (Dennett, 1991a, p. 199 – nossa tradução). Com efeito, os memes são uma espécie de simbioses com características mutualistas, ou seja, proporciona vantagens ao hospedeiro (cf. Dennett, 1998, p. 04). Que vantagens seriam essas? Em uma palavra: consciência.

A teoria memética, de acordo com Dennett (1991a; 1995a), emerge em detrimento da ideia da ilusão do significador central<sup>69</sup>. A desconstrução do teatro cartesiano depende da edificação de uma teoria mais realista que abarque de maneira mais científica possível a origem real dos atos de fala – atos que somos habitualmente inclinados a atribuir a um “eu”, um fantasma na máquina que realiza declarações. Nesse sentido, se isso for possível, e o autor está convicto que é, veremos “o que acontece com o mito habilitador da heterofenomenologia quando as complexidades da produção da linguagem recebem o devido valor” (Dennett, 1991a, p. 228 – nossa tradução).

Estamos certos de que quando falamos algo, realmente dizemos o que queremos dizer. Apesar de realizarmos uma experiência consciente em primeira pessoa – que nos proporciona um caráter privado dessa experiência – posso compartilhar um quinhão dessa experiência da maneira que melhor me aprouver. Por conseguinte, em uma escolha deliberada, tenho a prerrogativa de descrever ou não a minha experiência, seja ela passada ou atual. O processo de compartilhamento do meu estado mental necessita de que se articule frases que sejam meticulosamente associadas àquilo que pretendo relatar. Nesse sentido, há uma seleção das palavras que devem ser articuladas para expressar maior coerência em relação ao que quero relatar, observando as que possuem maior significado, que têm maior poder de descrição do

---

<sup>67</sup> De acordo com Dennett (1991a; 1991b), trata-se da concepção segundo a qual defende que os neurônios, em sua individualidade ignorante, proporcionam em conjunto a atividade consciente.

<sup>68</sup> O conceito de Memes foi desenvolvido por Richard Dawkins em sua obra *O gene egoísta*, de 1976, segundo a qual o autor apresenta a ideia de evolução cultural através do meme. Dawkins cunhou esse termo para descrever unidades de transmissão cultural: uma unidade de imitação. O termo é um cognato da palavra *Mimeme* (uma raiz grega adequada), mas Dawkins queria um monossílabo que soasse parecido com “gene”, algo que fosse relacionado a ‘memória’, dando um uso específico à palavra francesa meme (cf. Dennett, 1991a, p. 202).

<sup>69</sup> Tese dualista e, de certo modo, do materialismo reducionista.

meu caso particular. Ao entender que alcancei o nível de expressão mais satisfatório, mais explícito, torno-o público. Destarte, considerando o relato da minha experiência introspectiva, um outro indivíduo pode tomar conhecimento da minha experiência consciente. Mas como adquirimos nosso arcabouço argumentativo? Ele é inato? Ou é extraído da experiência? Se é extraído, como se dá esse aprendizado?

Segundo Dennett (1991a), é a nossa cultura que fornece nosso vocabulário. Há uma transmissão de caráter viral das palavras e frases proporcionadas pelos memes. Nesse sentido, palavras e frases são na verdade os aspectos fenotípicos mais proeminentes dos memes que nos “contaminam”, proporcionando a sua replicação, ou seja, um sistema de linguagem se configura como um ambiente perfeito no qual os “burocratas supervisores abdicaram parcialmente, cedendo uma grande medida de controle para as próprias palavras, que, na verdade lutam entre si por uma chance de serem os holofotes da expressão pública” (Dennett, 1991a, p. 243).

Portanto, a memética é uma concepção de caráter epidemiológica, sendo aplicável os mesmos modelos desse campo. Nosso cérebro é submetido a um estado virótico e isto é facilmente percebido. Do mesmo modo que somos invadidos por um sem número de vírus, somos também povoados por memes. Essa constatação pode ser verificada em circunstâncias em que “arquivamos” em nossa memória um refrão de uma música – ou uma palavra que não lembramos o significado – que fica “martelando” nossa consciência. Sem embargos, somos feitos de uma grande rede complexa de memes, os memplexes, eles “são as religiões, cujas expressões e orações, mesmo que fragmentadas, habitam o cérebro de quase todas as pessoas” (Teixeira, 2008, p. 24).

Comumente somos atraídos por determinados termos por gostarmos muito mais do modo como soam e damos pouca relevância ao seu real significado. A emergência de gírias, por exemplo, condiciona os membros de um grupo a adesão muitas vezes involuntária. Invariavelmente somos seduzidos por sentenças inteiras, para ampliar os níveis de agregação, não por seu valor descritivo ou atenderem especificações proposicionais preestabelecidas, mas pela natureza agradável em que ressoam em nossos ouvidos ou são pronunciadas (cf. Dennett, 1991a, p. 243).

Conforme Dennett (1991a; 1995a; 1998), podemos conceber melodias, frases de efeito, indumentárias, maneiras de fazer casas ou construir ferramentas como exemplos típicos de memes. Os genes se propagam em seu reservatório, de corpo em corpo, através de uma “contaminação” via espermatozoides ou óvulos. De modo similar, os memes também se replicam, saltam de cérebro para cérebro, via um processo que, num sentido amplo, pode ser denominado de imitação (cf. Dennett, 1991a, p. 202). Sem exagero, podemos afirmar que se

algum cientista desenvolve uma boa ideia, expressa em uma descoberta ou teoria, procura compartilhá-la com colegas e estudantes. Promove palestras sobre sua descoberta e escreve artigos abordando o tema. Se houver uma aceitação, podemos afirmar com idoneidade que será propagada, será repassada de cérebro para cérebro.

Em Dennett (1991a; 1995a), não devemos apenas estabelecer uma analogia do meme com a evolução biológica (genética) – não aplicar simplesmente uma metáfora a partir de uma linguagem evolucionária –, mas sim interpretar a ideia de evolução memética literalmente, ou seja, como um fenômeno que observa rigorosamente as leis de seleção natural. Ele argumenta que a teoria da evolução por seleção natural é de caráter de neutralidade no que concerne à distinção entre memes e genes – são tipos diferentes de replicadores que evoluem em campos distintos e em taxas nada comuns. Do mesmo modo que não seria possível existirem genes para animais sem que as plantas proporcionassem as condições necessárias (oferecendo uma atmosfera rica em oxigênio e fornecimento de nutrientes), o desenvolvimento dos memes não poderia ser iniciado até que fosse evoluída, dentre tantos animais, uma espécie particular com um cérebro que pudesse abrigar os memes, que pudesse nutrir hábitos de comunicação – os meios necessários para transmissão dos memes (cf. Dennett, 1991a, p. 202).

Destarte, devemos assimilar a consciência humana como derivada da seleção natural, mas fundamentalmente da evolução cultural. Nesse sentido, “a melhor maneira de ver a contribuição dos memes para a criação de nossas mentes é seguir de perto os passos padronizados do pensamento evolutivo” (Dennett, 1991a, p. 203 – nossa tradução). A esse respeito, Dennett (1991a, p. 208 – grifo nosso) assevera – e nos explica melhor:

Todos os três meios – evolução genética, plasticidade fenotípica e evolução memética – contribuíram para o *design* da consciência humana, cada um por sua vez, e em taxas crescentes de velocidade. Em comparação com a plasticidade fenotípica, que já existe há milhões de anos, a evolução memética significativa é um fenômeno extremamente recente, tornando-se uma força poderosa apenas nos últimos cem mil anos e explodindo com o desenvolvimento da civilização há menos de dez mil anos. É restrito a uma espécie, o *Homo sapiens*.

Nossos cérebros, de acordo com Dennett (1991b; 1995a), se configurariam como um tipo especial de ninho de memes que proporciona a mente humana. Tais entidades proporcionam também o aumento das possibilidades de replicação e transmissão entre si. Um dos exemplos dos reforços dos processos que permitem a implantação de memes seria a educação (cf. Dennett, 1991a, p. 207). Assim, devemos conceber que a “língua e a cultura fazem o papel de um *software* do cérebro a partir de cujos elementos são elaboradas as diferentes narrativas da máquina joyceana” (Teixeira, 2008, p. 24 – grifo nosso). O meme seria a “chave para entender como podemos ser não apenas os guardiões e transmissores de cultura,

mas também entidades culturais” (Dennett, 1998, p. 03 – nossa tradução). Assim, como vimos, os memes são uma espécie de simbionte com características mutualistas, ou seja, proporcionam vantagens ao hospedeiro. Essa vantagem é calcada pela evolução cultural, já que os fenômenos culturais são perfeitamente acolhidos pelo modelo de seleção artificial darwiniano, ou seja, são fenômenos linguísticos, sistemas arquitetados pela evolução – um processo segundo o qual a intervenção humana premeditada é mínima (cf. Dennett, 1998, p. 12).

A linguagem tem como principal atribuição a expressão do pensamento. As perguntas que cabem aqui são as seguintes: como podemos descrever, rigorosamente, em termos gerais algo que pressupõe a sua existência em uma introspecção? Se procuro descrever a mente com o método puramente intuitivo, deparo-me com o problema da subjetividade<sup>70</sup>, ou seja, como posso ter certeza de que aquilo que percebo na minha introspecção, os outros percebem da mesma forma? Se utilizar exclusivamente o método contra-intuitivo, configura-se a questão da impossibilidade da generalização subjetiva dos dados empíricos. Como posso garantir que o tom de azul que vejo agora na parede do meu quarto é a mesma tonalidade que minha esposa ou meu filho veem? Como posso estabelecer termos gerais para os *qualia*<sup>71</sup> – as experiências subjetivas e introspectivas de um indivíduo? Dennett (1991a; 1995a; 1998), acredita firmemente que com a teoria memética poderá dispensar essas questões, as resolvendo de modo definitivo.

De acordo com o filósofo, as experiências mentais existem virtualmente. Os memes seriam pequenos aplicativos que permitem uma representação dos estados mentais e da experiência qualitativa. Ao abordar a teoria memética como uma das subestruturas da teoria da consciência dennettiana é possível esclarecer que o autor aplica a tese do darwinismo universal, ao assimilar a evolução da cultura como um fenômeno proporcionado pela transmissão de um tipo especial de vírus. A esse respeito Dennett (1991a, p. 184 – nossa tradução) ressalta que:

de uma forma ou de outra, o cérebro plástico é capaz de se reorganizar de forma adaptativa em resposta às novidades particulares encontradas no ambiente do organismo, e o processo pelo qual o cérebro faz isso é quase certamente um processo mecânico fortemente análogo à seleção natural. Este é o primeiro novo meio de evolução: fixação do design pós-natal em cérebros individuais. Os candidatos à seleção são várias estruturas cerebrais que controlam ou influenciam os comportamentos, e a seleção é realizada por um ou outro processo de eliminação mecânica que está geneticamente instalado no sistema nervoso.

A distinção entre os teóricos da I.A. e os fisicalistas funcionalistas reside no fato de que os primeiros não são compromissados com uma ontologia particular da mente. Contudo, o fato

---

<sup>70</sup> O denominado problema das outras mentes (cf. Churchland, 2004 p. 16-17).

<sup>71</sup> Estados mentais podem apresentar dois tipos de conteúdo: (1) o conteúdo representacional/intencional e o (2) conteúdo qualitativo, os *qualia* – embora Dennett não estabeleça com clareza essa distinção.

do surgimento dos computadores nos forneceu uma prova de existência de influência inegável: existem mecanismos – mecanismos brutos e pouco misteriosos operando de acordo com princípios físicos rotineiramente bem compreendidos – que têm muitas das competências até agora atribuídas apenas à mente humana. A esta altura o leitor já deve ter percebido que, para Dennett, o físico não deve ser confundido com o material.

Dennett (1999) acredita firmemente, devido à similaridade da estrutura cerebral com sistemas computacionais, em uma explicação naturalista (darwinismo) e mecanicista (ciência computacional) da consciência. Para ele, como já dissemos, as experiências mentais existem virtualmente. A partir da década de 1990, ele procura moldar as suas ideias anteriores numa perspectiva evolucionista assimilando a evolução por seleção natural como um processo algorítmico. Em sua teoria memética percebemos que o cérebro está a todo momento criando narrativas, versões sobre percepções, sensações, emoções e sentimentos. Nesse ponto da teoria dennettiana, em substituição ao teatro cartesiano, nos é apresentado a noção de mente como oriunda de rascunhos múltiplos (*Multiple Drafts*) formados por memes. Dennett (1991a, p. 128 – grifo nosso) realiza uma observação da sua importância na seguinte passagem:

uma vez que não existe tal teatro, não existe tal público [...] o modelo *Multiple Drafts* afirma que o cérebro não se preocupa em “construir” nenhuma representação que se dê ao trabalho de “preencher” as lacunas. Isso seria uma perda de tempo e (digamos?) pintura. O julgamento já está feito, então o cérebro pode continuar com outras tarefas.

No cérebro, homúnculos – circuitos especializados – trabalham em paralelo criando fragmentos de narrativas e vão se digladiando o tempo todo entre si, em busca da luz da consciência e da publicação. O autor, fundamentado em um darwinismo neural – onde esboça a ideia de que os neurônios, em sua individualidade ignorante, proporcionam em conjuntos a atividade consciente – e em uma psicologia computacional, afirma que a mente seria um *software* que faz *downloads* (memes/linguagem e cultura) por intermédio do cérebro (o *hardware*) que permite *inputs* e *outputs*. Essa concepção de mente sobrepuja tanto a ideia de mente como algo que possui existência própria, quanto o entendimento de que ela é um produto do cérebro (cf. Dennett, 1991a, p. 216).

Dissemos com Dennett (1998) que os memes são itens da cultura que têm aptidões e uma história evolutiva tanto quanto um *gene*, itens culturais análogos a um vírus. Um vírus é um ácido nucleico com “atitude”. Ele não está vivo, mas evolui por seleção natural, ou seja, a evolução natural não ocorre apenas entre seres vivos – ocorre em qualquer sistema de entidades autoreplicativa que podem ter diferentes poderes de replicação. Os memes promovem mutações no hospedeiro e no ambiente. Ele é uma estrutura de dados composta de informação com “atitude” – com o termo atitude deve-se entender que alguns memes se replicam melhor que

outros. Mas, os memes existem? Palavras são memes que podem ser pronunciados. Acreditamos em palavras? Pensamos que elas existem? Se nossa resposta for sim, então acreditamos em memes. Afinal, o que são palavras?

Segundo Dennett (2009a), um dos aspectos da evolução que se tornou muito claro é que para acontecer a cópia precisa ser de alta fidelidade, mas não perfeita. Se não for de alta fidelidade todas as invocações desaparecerão rapidamente. Se for perfeita, não haverá mutação e então não haverá inovação. Os memes são algo que evolui nas línguas e é a chave para o poder que a língua tem de ser copiada e se replicar com índice muito baixo de mutação – que é exatamente o que precisamos. Essas normas de correção lhes dão uma espécie de alfabeto de possibilidades que digitaliza o espaço e possibilita a replicação à alta fidelidade. Os memes são vírus de *software*.

Portanto, para entender o que é um meme, de acordo com Dennett (2006), é necessário regular nosso pensamento sobre computação e *software*. Para entender nossa mente como um *software*, precisamos estabelecer uma analogia da linguagem com a ciência da computação. Palavras são *máquinas virtuais*, são instruções complexas de comportamento – há uma combinação de significado, semântica, som, fonologia e de regras sintáticas. Quem as projeta? Ninguém – evoluem por seleção natural. Mas, não por seleção genética (não existe um gene para o alemão, por exemplo). O idioma que sabemos, aprendemos por transmissão cultural. As próprias palavras são pequenas máquinas virtuais, pequenos aplicativos, e a competência que temos deriva do *software* (projetado pela evolução cultural) instalado em nosso cérebro.

Nesse sentido, Dennett (1998) dá, também, uma explicação genética para a evolução da língua. Ele afirma que os cérebros se tornaram mais competentes em línguas, mas a língua também se tornou mais eficaz para se instalar em um cérebro, se tornando assim um processo coevolutivo – sendo que a língua evolui mais rápido que o cérebro. Como instalamos máquinas virtuais? Em uma atividade dinâmica que envolve repetição, ou seja, por processos infantis de aquisição de língua. Assim, os nossos cérebros são processadores de palavras mais eficazes do que quando os nossos ancestrais começaram a falar.

Conforme Dennett (1991a), em detrimento ao modelo burocrático cartesiano – tese que, segundo o autor, nenhum fisicalista defende, mas que todos utilizam como pressuposto teórico –, devemos considerar um *pandemonium* de narrativas<sup>72</sup>, cada uma buscando os holofotes da

---

<sup>72</sup> Refutação do teatro cartesiano usando o exemplo de percepção inconsciente e ação inteligente (é como se outra pessoa estivesse dirigindo).

expressão pública. Se há alguma dúvida das concepções de Dennett (1991a, p. 238 – nossa tradução), doravante poderão ser sanadas através de suas palavras:

Podemos supor que tudo isso acontece em gerações rápidas de processamento paralelo "perdulário", com hordas de demônios anônimos e suas construções esperançosas que nunca veem a luz do dia – seja como opções que são conscientemente consideradas e rejeitadas, ou como atos de fala finalmente executados para estranhos ouvirem.

De acordo com Dennett (1991a; 1995a; 1998), as palavras que se rivalizam (os demônios-palavras) paralelamente são os questionadores e competidores, enquanto os demônios conteudistas são respondentes e júri, ou seja, os significados são oriundos de um processo um tanto quanto evolutivo de *design* de atos de fala<sup>73</sup>, permeado de colaboração – serial e paralelo – de uma variedade de subsistemas segundo os quais, por si só, são incapazes de proporcionar a execução (ou determinação) de um ato de fala (cf. Dennett, 1991a, p. 238). Esse é o modo pelo qual o autor soluciona o problema que teríamos de “encontrar uma maneira pela qual as fontes de conteúdo possam influenciar ou restringir as energias criativas dos demônios-palavra sem ditá-los” (*Ibidem*). O comentário de Dennett (1991a, p. 240 – nossa tradução) é esclarecedor:

No outro extremo estão os modelos em que palavras e frases do léxico, juntamente com seus sons, significados e associações, se chocam com construções gramaticais em um pandemônio, todas "tentando" fazer parte da mensagem, e algumas delas, portanto, dar uma contribuição substancial às próprias intenções comunicativas que ainda menos delas acabam por executar [...] não há uma fonte de significado, mas muitas fontes mutantes, oportunisticamente desenvolvidas a partir da busca pelas palavras certas. Em vez de um determinado conteúdo em um local funcional específico, esperando para ser traduzido por sub-rotinas, há uma mentalidade ainda não completamente determinada distribuída pelo cérebro e restringindo um processo de composição que, com o passar do tempo, pode realmente retroalimentar, fazer ajustes ou revisões, determinando ainda mais a tarefa expressiva que colocou o processo de composição em movimento em primeiro lugar.

De acordo com esse paradigma, o modelo pandemônio, o controle do processo de enunciados não é delegado a um administrador central, não há um planejamento, mas é caracterizado pelo oportunismo das fontes múltiplas que estabelecem as decisões de *design* que proporcionam o enunciado que será publicado. Não é possível delinear uma zona que limite “as ordens de marcha do conteúdo fluindo de dentro e as sugestões voluntárias para implementações apresentadas pelos demônios-palavra” (Dennett, 1991a, p. 241 – nossa tradução). Esse modelo, portanto, alvitra que para conservar o poder criativo daquele que expressa o pensamento é necessário se abster da ideia de que o processo de consciência se inicia

---

<sup>73</sup> O termo aqui utilizado não intenciona uma associação com a concepção de John Austin e sua teoria de atos de fala.

com um pensamento determinado, previamente delineado. A esse respeito, Dennett (1991a, p. 251 – grifo nosso) esclarece:

A patologia, seja a tensão temporária induzida por experimentos inteligentes ou os colapsos mais permanentes causados por doenças, ou danos mecânicos ao cérebro, fornece uma abundância de pistas sobre como o mecanismo está organizado. Esses fenômenos me sugerem que nossa segunda caricatura, *pandemonium*, está mais perto da verdade do que um modelo burocrático mais digno estaria, mas isso ainda precisa ser submetido ao teste empírico adequado. Não estou afirmando que seja impossível para um modelo amplamente burocrático fazer justiça a essas patologias, mas apenas que elas não parecem ser as falhas naturais de tal sistema.

O *pandemonium* é um modelo cognitivo que se configura como fundamental na teoria dennettiana. Considerando sua aplicação original<sup>74</sup>, esse modelo seria, por exemplo, um *software* com um sistema de vinte e três demônios segundo os quais cada um teria a competência de reconhecer uma letra do alfabeto. Uma sequência de palavras é apresentada aos demônios e cada um escolhe a sua letra programada. Sendo “a” a primeira letra, o demônio correspondente acusará identificação – o demônio correspondente à “o” poderia duvidar e os referentes aos “b”, “c” ou “d” não irão admitir terem reconhecido um “a”. Aqueles que realizaram a identificação irão promover uma prestação de contas com um demônio-chefe encarregado de fazer um levantamento dos dados e verificar quem é o vencedor. A eficácia desse processo se dará se o sistema for submetido a um número considerável de repetições – algo em torno de dez ou quinze mil vezes, por exemplo –, ou seja, só assim o sistema será capaz de realizar reconhecimento com maior precisão.

De modo similar, segundo Dennett (1991b), tais processos ocorrem em nosso cérebro. Nossos estados mentais seriam determinados por uma competição de milhares de demônios (os pequenos agentes). Alguns se sobressaem em relação a outros, mas essa predominância é temporária. Não há um caminho simples que se estabeleça entre um estímulo e nossa percepção e não é preciso identificar um local ou momento específico segundo o qual seja inserido no campo fenomênico, ou seja, onde ganhe o estatuto de consciente. Segundo o filósofo, o que há, na verdade, é uma multiplicidade de canais paralelos.

Embora o modelo *pandemonium* explique, com relativa plausibilidade, como podemos dispensar a tese que enaltece a necessidade de um burocrata, ainda nos falta descrever o que é

---

<sup>74</sup> Na década de 1950, o pesquisador do MIT (Massachusetts Institute of Technology) Oliver Selfridge trabalhava em um *software* que teria a capacidade de proporcionar aprendizado a um computador – essa ambição era extremamente revolucionária para época, pois, nesse período, os computadores se resumiam a processamento de informações de forma passiva, a partir de um conjunto de regras de cálculo. Selfridge utiliza o termo *pandemonium* (todos os demônios) para caracterizar seu programa, qual seja, a capacidade de reconhecer e identificar padrões inconstantes ou mal definidos. Esse programa de computador revolucionou a ciência da computação, trazendo o entendimento de que a inteligência se dá de forma distribuída e não centralizada, ou seja, o desenvolvimento da inteligência artificial não se dá por meio de um programa único centralizador, mas por um conjunto considerável de miniprogramas (cf. Teixeira, 2008, p. 22).

essa estrutura formada por demônios-palavra. Trata-se do já citado modelo de consciência de rascunhos múltiplos (*multiple drafts*) – a primeira versão de substituição efetiva do teatro cartesiano – é um “fluxo de conteúdo e apenas mais parecido com uma narrativa por causa de sua multiplicidade; em qualquer ponto do tempo, existem múltiplos ‘rascunhos’ de fragmentos narrativos em vários estágios de edição em vários lugares do cérebro” (Dennett, 1991a, p. 113). Segundo essa perspectiva, são desenvolvidos no cérebro processos concomitantes de interpretação e desenvolvimentos de entradas sensoriais em várias trilhas diferentes proporcionando as percepções e o pensamento (a atividade mental em geral). Esses conteúdos que são reunidos no cérebro estão constantemente sob revisão editorial. Dennett (1991a, p. 112 – nossa tradução) continua em sua análise, acerca dos rascunhos múltiplos, argumentando que:

esses processos editoriais ocorrem em grandes frações de segundo, durante as quais podem ocorrer várias adições, incorporações, emendas e substituições de conteúdo, em várias ordens. Não experimentamos diretamente o que acontece em nossas retinas, em nossos ouvidos, na superfície de nossa pele. O que realmente experimentamos é produto de muitos processos de interpretação – processos editoriais, na verdade.

Esse modelo, segundo o autor, nos permite evitar o erro de considerar que há uma única interpretação, a realizada pelo burocrata. Assim, estamos diante de “uma nova visão da consciência, que envolve uma grande reforma (mas não uma revolução radical) em nossas maneiras de pensar sobre o cérebro” (Dennett, 1991a, p. 114 – nossa tradução). Essa é a implicação fundamental do modelo de rascunhos múltiplos, pois, tal modelo deixa explícito que, ao ser desenvolvida, uma discriminação não carece ser reelaborada já que “o cérebro apenas se ajusta à conclusão a que se chega, disponibilizando a nova interpretação da informação para a modulação do comportamento subsequente” (Dennett, 1991a, p. 127 – nossa tradução).

Ao ser estabelecida, essa discriminação se torna o fundamento de um determinado comportamento, seja ele a modulação de um estado informativo interno (um reconhecimento de um padrão, por exemplo) que rotulamos de consciência, um comentário ou até mesmo apertar um botão. Ao discriminar uma foto de um gato, por exemplo, o cérebro tende a criar um “conjunto perceptível” – que facilita a identificação temporária de gatos em outras fotos – ou proporcionar a possível ativação de domínio semântico específico, permitindo uma brevidade interpretativa segundo a qual seríamos mais inclinados a ler a palavra “gato” como o animal, ao invés de gambiarra para roubar energia, por exemplo (cf. Dennett, 1991a, p. 135). Esse processo se dá em um tempo absurdamente curto (em centenas de milissegundos), mas permite a ocorrência – em várias ordens – de adições, emendas, substituições de conteúdo e incorporações. Tudo isso, no decorrer do tempo, faz emergir uma espécie de sequência ou fluxo

narrativo que, erroneamente, pode ser entendido como um sujeito, mas que, na verdade, é a contínua edição de muitos processos distribuídos no cérebro de modo perpétuo. Nesse sentido, no cérebro emergem conteúdos que são revisados e ajudam na interpretação de outros conteúdos ou na modulação de comportamentos, sejam eles verbais ou não, e essa atividade lastreia a memória onde podem decair, ser incorporados ou substituídos (total ou parcialmente) por novos conteúdos. Há, portanto, um emaranhado de conteúdos vislumbrando a publicidade distribuído no cérebro (cf. Dennett, 1991, p. 135).

De acordo com Dennett (1991a), alguns conteúdos se caracterizam por ser efêmeros – alguns deles realizarão sua breve contribuição e serão extintos em seguida, enquanto outros sequer contribuirão – e uma outra parcela insistirá em desempenhar uma multiplicidade de delegações na modulação adicional do comportamento e do estado interno. Outros ainda perseveram em manter sua estadia processual sob a forma de comunicado de imprensa publicados na forma de comportamento verbal. Analisar esse fluxo de narrativas em intervalos variados, segundo o filósofo, culmina em efeitos distintos, uma variação de narrativas – versões exclusivas de uma porção de fluxo de consciência. Adiado a análise por longos períodos, corre-se o risco da não obtenção de narrativas, por outro lado, se anteciparmos a análise “pode-se coletar dados sobre quão cedo uma discriminação específica é alcançada no fluxo, mas ao custo de interromper a progressão normal do fluxo” (Dennett, 1991a, p. 136 – nossa tradução). A esta altura o leitor já percebeu que nenhum desses rascunhos – fragmentos de narrativas – pode ser o gerenciador da consciência, pois os “rascunhos que foram compostos em momentos e lugares específicos do cérebro foram posteriormente retirados de circulação, substituídos por versões revisadas, mas nenhum deles pode ser apontado como definitivo do conteúdo da consciência” (Dennett, 1991a, p. 142 – nossa tradução).

Assim que esses “órgãos especializados de evolução independente” são reunidos em prol do mesmo objetivo – e essa união proporciona o aprimoramento de seus poderes –, esse *software* cerebral realiza uma miraculosa façanha “política” interna: permite o surgimento de uma espécie de líder virtual do grupo, um homúnculo que não possui a prerrogativa de um poder gerencial de longo prazo. Quem seriam os delegados? A princípio um grupo e depois outro, numa sequência ordenada por “bons meta-hábitos que tendem a incorporar sequências coerentes e intencionais, em vez de uma interminável tomada de poder desordenada” (Dennett, 1991a, p. 228 – nossa tradução).

Não obstante, o fisicalismo ingênuo, que na verdade é uma versão do dualismo da substância, ou seja, a doutrina que todos negam, mas em cujos termos todos estão inclinados a pensar, oferece a seguinte estrutura: há um tráfego de informações pelo cérebro que é

processado por uma variedade de mecanismos em vários lugares. O testemunho de nossas intuições alvitra que os fluxos de consciência se resumem a eventos sequenciados. Além disso, a cada momento determinado subitamente, esses elementos sequenciais são submetidos a uma classificação parcial que notifica como “já ocorrido” e “ainda não ocorrido” na consciência. No entanto, se de fato a emergência da consciência se dá nesses parâmetros, devemos considerar que a circulação, no cérebro, dos “transportadores de conteúdos” carregados determinam a ordem em que passam por determinado ponto, e essa ordem determina a chegada no teatro da consciência, onde efetivamente se tornam conscientes (cf. Dennett, 1991a, p. 144).

Nesse sentido, segundo Dennett (1991a), o ponto de convergência de teorias calcadas no cartesianismo, ou seja, de caráter burocráticas, é o fato de que estabelecem uma distribuição dos homúnculos segundo uma predefinição de hierarquias. Há, portanto, uma estrita regulamentação que organiza a competição entre eles. Os homúnculos<sup>75</sup> da teoria pandemônio, por outro lado, “postulam muita duplicação de esforços, desperdício de movimento, interferência, períodos de caos e vagabundos sem descrição de trabalho fixo” (Dennett, 1991a, p. 261). Os homúnculos em questão possuem competências circunscritas individuais e as teorias consideram essas unidades para elaborar uma descrição de como atribuições maiores podem ser executadas a partir de organizações de unidades que realizam pequenas funções. Toda teoria funcionalista – da mais rigorosa à mais artificial de caráter metafísica – partem desse pressuposto. Assim, toda versão do funcionalismo é em si uma abordagem calcada em uma noção de funcionalismo homuncular. Sobre os “demônios especialistas”, Dennett (1991a, p. 263 – nossa tradução) oferece mais detalhes:

Eles [...] são frequentemente alistados oportunisticamente em novos papéis, para os quais seus talentos nativos mais ou menos os adaptam. O resultado não é confusão apenas porque as tendências que se impõem a toda essa atividade são, elas mesmas, produto do design. Parte desse design é inato e compartilhado com outros animais. Mas é aumentado, e às vezes até superado em importância, por micro-hábitos de pensamento que são desenvolvidos no indivíduo, em parte resultados idiossincráticos da autoexploração e em parte os dons pré-definidos da cultura. Milhares de memes, principalmente transmitidos pela linguagem, mas também por “imagens” sem palavras e outras estruturas de dados, fixam residência em um cérebro individual, moldando suas tendências e, assim, transformando-o em uma mente.

A máquina joyceana<sup>76</sup> dennettiana caracteriza a consciência fenomênica (o pensamento) fundamentalmente como uma espécie de reordenamento de *inputs* associado a uma narrativa

---

<sup>75</sup> Segundo Dennett (1991a), faz mais sentido denominar essas unidades de homúnculos (agentes ou demônios) em sua teoria do que simplesmente de unidades.

<sup>76</sup> Uma espécie de máquina virtual (a consciência para Dennett) com a atribuição de selecionar as melhores narrativas do *pandemonium* que lutam entre si por seus “15 minutos de fama”. Nesse sentido, a máquina joyceana – uma alusão ao estilo literário do escritor irlandês James Joyce (1882-1941) caracterizado por uma narrativa em fluxo – seria um enorme complexo de memes que tem sua potência elevada graças aos “poderes subjacentes desse

efêmera. O experimento realizado com o fenômeno *phi* – que proporcionou o surgimento da psicologia da Gestalt – nos ajuda a compreender a proposta dennettiana:

O fenômeno *phi* é um caso de movimento aparente, estudado, inicialmente, pelo psicólogo Max Wertheimer: duas ou mais luzes piscam sucessiva e rapidamente, causando a impressão de movimento aparente. Dennett nos apresenta uma pequena variação desse experimento, inserindo, em um dado momento, uma luz de cor diferente. Neste caso, o sujeito percebe o movimento aparente como antes, mas vê um ponto luminoso que se move e muda a cor gradualmente. A pergunta que os psicólogos têm feito é saber como é possível que um evento físico possa interferir em uma percepção anterior a ele. Ou, em outras palavras, saber como é possível que a mente detecte um estímulo antes que o cérebro tenha tempo de processá-lo. Dennett aponta para dois tipos possíveis de explicação para o fenômeno *phi*, de acordo com seu modelo de consciência: a reconstrução orwelliana [...] e a stalinista (Teixeira, 2008, p. 23 – grifo nosso).

De acordo com Dennett (1991a), a abordagem orwelliana<sup>77</sup> defende a tese que o sujeito capta o primeiro estímulo integralmente, porém, essa primeira interpretação é preterida a partir de uma edição, dando origem a uma segunda versão que se tornará definitiva. Por outro lado, a abordagem stalinista<sup>78</sup> pressupõe a tese de que os estímulos recebem uma edição antes mesmo de se tornarem conscientes, não há uma versão integral consciente, ou seja, são editados na sua origem, prevalecendo desta forma a segunda interpretação desde o princípio. O trabalho do cérebro é de edição de narrativas coerentes de caráter efêmero e às vezes muito fragmentárias. Nesse sentido, essas narrativas dispensam a necessidade de sua emergência em tempo real – não precisa ocorrer “no tempo físico de recepção de estímulos e de condução dos impulsos nervosos” (Teixeira, 2008, p. 23).

O meme é um tipo de vírus que é transmitido pela linguagem. A inteligência consciente é característica do ser humano justamente devido ao fato do cérebro ser estruturado e esculpido pela evolução natural para “baixar” esses dados da cultura (cf. Dennett, 1991a, p. 210). Do mesmo modo que concebemos o surgimento da vida a partir das atividades de elementos físicos não vivos, a consciência deve ser entendida como atividades de estruturas não conscientes.

Dennett (1998) ironiza os críticos do conceito de memes aplicando a definição da lei do esturjão cujo entendimento é de que mais de 95% desses argumentos não possuem fundamento e não são passíveis de atenção, pois, são calcados em um mal-entendido – não devemos considerar o modelo lamarckiano para acomodar a estrutura que concebe a evolução cultural, portanto, a consciência. Lograremos maior êxito, se considerarmos o modelo darwiniano que,

---

*hardware* orgânico sobre o qual ela se instala, numa espécie curiosa de reutilização de um órgão que existia com outras finalidades” (Teixeira, 2008, p. 24 – grifo nosso).

<sup>77</sup> Uma referência ao livro 1984, de George Orwell, no qual os governantes reescreviam a história da maneira que melhor lhe apropsessem.

<sup>78</sup> Denominado assim por Dennett (1991a) para estabelecer uma analogia à censura da imprensa na antiga União Soviética, instaurada por Josef Stálin (1878-1953).

a rigor, é uma espécie de seleção artificial parcialmente consciente de fenômenos linguísticos. Entrementes, segundo o filósofo, restam 5% das críticas dignas de atenção. Dentre elas, iremos destacar, no capítulo III, a desenvolvida por John Searle.

## 2.4 A formulação da teoria dos Sistemas Intencionais<sup>79</sup>

Dennett (1991a) defende que devemos ir além da fenomenologia “pura” – a mera aceitação dos fenômenos da percepção –, não devemos nos deixar influenciar por simples apreensões introspectivas. Se quisermos realmente entender o que são estados mentais, devemos observar os aspectos subjacentes a esses fenômenos. Nesse sentido, essa concepção de fenomenologia comete o erro de promover a crença de que a experiência introspectiva se dá rigorosamente do modo que o indivíduo a percebe. (cf. Dennett, 1991a, p. 67), ou seja, existe um excesso de confiança nas experiências internas.

De acordo com Dennett (1991a; 2006), os psicólogos cognitivos e os estudiosos da inteligência artificial (I.A.) ignoram a fenomenologia pura por não a considerar funcional – o que não significa que deixam de nutrir um fascínio por ela. Procurando esclarecer os eventos fenomenológicos, argumentam que o uso da imagem interna não pode ser o principal pilar da linguagem para delinear a compreensão (pensamento). Por ora, vejamos o que o filósofo diz a esse respeito:

A natureza quase visual da fenomenologia da compreensão foi quase totalmente ignorada por pesquisadores em ciências cognitivas, particularmente em inteligência artificial, que tentaram criar sistemas de computador para compreensão de linguagem. Por que deram as costas à fenomenologia? Provavelmente em grande parte por causa de sua convicção de que a fenomenologia, por mais real e fascinante que seja, não é funcional – uma roda que gira, mas não envolve nenhum mecanismo importante de compreensão (Dennett, 1991a, p. 56 – nossa tradução).

No entanto, de acordo com Dennett (1991a), é uma incorreção falar em fenomenologistas como especialistas sobre as coisas que habitam o mar misterioso da consciência. Diferentemente da abordagem dualista e materialista, “podemos seguir a prática recente e adotar esse termo (com f minúsculo) como um termo genérico para os vários itens da experiência consciente que precisam ser explicados” (Dennett, 1991a, p. 45 – nossa tradução). Todos aqueles que, desde Descartes, examinaram a consciência, partiram do pressuposto de

---

<sup>79</sup> Em um sentido amplo, o termo intencionalidade (*Intentional* – escrito com “T” – se aproxima do sentido husserliano) corresponde ao ato de transcendência da consciência em direção a um objeto, ou seja, uma canalização da consciência para um objeto ou fato. No entanto, em Dennett o termo intencionalidade é derivado do inglês *Intensional* (escrito com “S”), os estados mentais não são considerados “por sua natureza intencional, a intencionalidade é apenas um meio para adquirir informações referentes a determinados objetos, para que seja possível comportar-se de modo feliz; há portanto objetos em relação aos quais a atribuição intencional é inútil (os objetos físicos em geral porquanto inanimados), outros em relação aos quais é indispensável (os seres humanos)” (Abbagnano, 2012, p. 663).

uma análise em primeira pessoa. Nessa perspectiva, um exame de caráter introspectivo, fenomenológico, não pode render o fruto da verdade científica, pois, esbarra no problema metodológico – como posso realizar generalizações a partir da minha experiência mental particular? No entanto, a força e o vigor que são defendidas nossas experiências mentais são refletidas, erroneamente, na edificação da maioria das teorias da mente (cf. Dennett, 1991a, p. 67).

Por conseguinte, se faz necessário delinear um método descritivo claro e neutro, pois, segundo o autor, é “evidente que duas pessoas não utilizam termos linguísticos do mesmo modo e ainda assim todos querem se passar por especialistas” (Dennett, 1991a, p. 70 – nossa tradução). A perspectiva pluralista em primeira pessoa é um ninho de erros e “na história da psicologia [...] foi o crescente reconhecimento desse problema metodológico que levou à queda do introspeccionismo e à ascensão do behaviorismo” (*Ibidem*). Essa perspectiva fica mais clara quando Dennett (1991a, p. 71 – nossa tradução) comenta que:

Todas essas visões levam a uma conclusão injustificada ou outra. Mesmo que os eventos mentais não estejam entre os dados da ciência, isso não significa que não possamos estudá-los cientificamente. Os buracos negros e os genes não estão entre os dados da ciência, mas desenvolvemos boas teorias científicas sobre eles. O desafio é construir uma teoria dos eventos mentais, usando os dados que o método científico permite.

Dennett (1999) entende que ao serem rotuladas de teorias, e não de hipóteses, as abordagens dualista e materialista procuram se impor como leituras exatas de como o cérebro e a mente se relacionam. No entanto, essas teorias não passam de hipóteses devido à sua fragilidade descritiva. Decerto, o fato de que o autor critica duramente o dualismo cartesiano – até mesmo o dualismo naturalista de David Chalmers – não o coloca nas fileiras dos fisicalistas reducionistas, pois, sua teoria se configura como uma terceira alternativa a essas abordagens da mente. O que torna o funcionalismo homuncular desenvolvido por Dennett um diferencial? Para responder a essa questão, é preciso examinar o método por ele adotado.

Já vimos que, de acordo com Dennett (2006), uma teoria da mente não deve ter como pressuposto uma análise em primeira pessoa – como propôs Descartes e os fenomenologistas – , mas em terceira pessoa, ou seja, segundo o viés científico. Se há um excesso de confiança nas experiências internas, devemos adotar, para evitar esse tipo de erro, um método que esteja em consonância com os trabalhos científicos, ou seja, um tipo de metodologia que desconsidera o testemunho do sujeito no que diz respeito às experiências de seus próprios estados mentais, que assume uma postura de neutralidade construtiva (cf. Dennett, 1991a, p. 83).

Assim, em detrimento de uma abordagem calcada em uma psicologia popular – em que se recorre a um método de análise exclusivamente em primeira pessoa para realizar

interpretações –, aplica-se o já comentado método heterofenomenológico. Essa metodologia permite ao filósofo esboçar sua teoria dos sistemas intencionais. De forma bastante eloquente, Dennett (2009b, p. 339 – grifo nosso) contribui conosco:

A teoria dos sistemas intencionais é, em primeiro lugar, uma análise dos significados de tais termos *mentalistas* como acreditar, desejar, esperar, decidir e pretender, os termos da psicologia popular [...] que usamos para interpretar, explicar e prever o comportamento de outros seres humanos, alguns artefatos, como robôs e computadores, e de fato nós mesmos.

No entanto, esse método não se configura por uma negação sumária da totalidade de nossas experiências conscientes, não obstante devemos ter o cuidado de não sermos arrogantes ou ingênuos assumindo uma crença de que nossos estados mentais são autoevidentes (cf. Dennett, 1991a, p. 68). Essa é a postura da psicologia popular. Sua grande limitação é a descrição do testemunho da experiência em primeira pessoa, sem examinar, ou sequer considerar, os fundamentos que proporcionam suas causas e seus efeitos. Sobre o erro dos fenomenólogos, Dennett (1991a, p. 70 – nossa tradução) ressalta que:

Os fenomenólogos tendem a argumentar que, uma vez que não temos autoridade sobre as causas e efeitos fisiológicos de nossa fenomenologia, devemos ignorar tais causas e efeitos em nossa tentativa de fornecer uma descrição pura, neutra e pré-teórica do que encontramos “dado” no curso de nossa fenomenologia.

Conforme Dennett (2006), sua teoria dos sistemas intencionais tem o desígnio de promover uma revisão, um exame mais apurado dos significados dos termos mentalistas. Ela não só é a base da sua psicologia, mas também serve de pano de fundo para sua teoria da mente – em detrimento de uma abordagem calcada na psicologia popular, segundo a qual comumente recorreremos para realizar uma interpretação –, bem como para efetuar previsões comportamentais de entidades (seres humanos, robôs ou algum outro artefato) de caráter inteligente (cf. Dennett, 2009b, p. 339).

Para um maior entendimento do que venha a ser um sistema intencional é necessário examinar o seu contraste com duas posturas mais simples que recorreremos para realizar previsões, a saber: a *postura física* e a *postura de design*<sup>80</sup>. A primeira é aquela que a partir da constituição material é possível realizar uma previsão de comportamento. As previsões e explicações são calcadas nas leis irrevogáveis da natureza. Observe um recipiente com água, por exemplo, submetido a aquecimento ao nível do mar. O resultado desse experimento é conhecido. Ao atingir 100°C ela começa a evaporar. Assim, como a repetição dessa experiência leva sempre ao mesmo resultado, temos uma prerrogativa que nos permite realizar uma previsão

---

<sup>80</sup> Denominada por Dennett (1997) de postura de planejamento. Segundo o autor, são observadas também em “projetos” naturais como, por exemplo, o semear e colher (cf. Dennett, 1997, p. 26).

de um dos comportamentos da água frente a elevação de temperatura, pois, a causa subjacente a esse fenômeno é conhecida pela física. Analogamente podemos atribuir ou expandir esse comportamento a outros líquidos, ou seja, também possuem um ponto de ebulição. Com efeito, a postura física é “simplesmente o método laborioso padrão das ciências físicas, em que usamos o que sabemos sobre as leis da física e a constituição física das coisas em questão para conceber nossa previsão” (Dennett, 2009b, p. 339 – nossa tradução). A segunda é caracterizada por previsões realizadas a partir do conhecimento da finalidade do ente em questão. Considera-se que para atingir determinado objetivo o ente deve assumir determinados comportamentos que promovam o êxito de sua finalidade. São obviamente mais difíceis de realizar previsões se comparadas às de caráter de postura física, haja vista que temos que considerar hipóteses adicionais de comportamento – como por exemplo, o mal funcionamento (cf. Dennett, 1997, p. 26).

Considere um forno de micro-ondas. Sua estrutura material segue os aspectos necessários das leis da natureza (postura física), mas sua organização física nos permite fazer previsões de comportamento contingentes. A rigor, todo projeto tem uma finalidade e é a partir do conhecimento desse objetivo específico que realizamos nossas previsões – no caso do forno de micro-ondas, aquecer alimentos. Entrementes, o projeto pode falhar e destoar da finalidade segundo a qual foi desenvolvido (programamos, por exemplo, o aparelho para aquecer por dois minutos, mas ele o faz por dez minutos). Enquanto as previsões na postura física são muito mais exatas (se arremesso uma pedra para cima a lei da gravidade fará com que procure o centro da terra) e não admitem falhas (a pedra não “quebra” nem sua finalidade é quebrar, se por ventura se dividir ao meio esse segundo comportamento é, na verdade, uma obediência à outra lei da natureza). Dennett (2006, p. 35) prossegue em nos explicar:

O aspecto essencial da postura de projeto é que fazemos previsões apenas a partir do conhecimento ou das suposições sobre a constituição funcional do sistema, não importando sua constituição física ou as condições das partes internas de um objeto particular.

Por conseguinte, esse tipo de postura possui previsões de caráter mais irregulares, seus vaticínios são mais passíveis de erro em relação à postura física (cf. Dennett, 2009b, p. 340).

A postura intencional é um *Projeto Ótimo*<sup>81</sup> uma espécie de postura de planejamento (design) mais complexo. Uma postura em que a finalidade é realizada por um agente observando uma série de fatores que facilitam e determinam o momento exato de execução. Essa postura requer instruções previamente estabelecidas (cf. Dennett, 1997, p. 27). A rigor,

---

<sup>81</sup> Racionalidade objetiva – trata-se de uma série de pequenos objetivos organizados hierarquicamente em função de uma finalidade maior (cf. Dennett, 2006, p. 38).

realizamos uma antropomorfização<sup>82</sup> de entes a partir da sua análise comportamental. No exame da postura intencional, estamos diante do conceito basilar da psicologia dennettiana, pois, o filósofo a considera fundamental para desvendar os mistérios de qualquer tipo de mente. Embora não seja aplicável – ou irrelevante – um juízo de valor, a postura intencional cumpre bem o seu papel de realizar predições com base no comportamento do agente. Dito de outro modo, independente do motivo que leve o agente a buscar a concretização de sua finalidade, o que o faz agir como age, o enquadra como manifestante de postura intencional. Em termos concretos, isso significa que “os sistemas intencionais são, por definição, todas e apenas aquelas entidades cujo comportamento é predizível ou explicável a partir da postura intencional” (Dennett, 1997, p. 30 – nossa tradução).

De acordo com Dennett (2009b), é falacioso argumentar que há uma subclassificação de sistemas intencionais<sup>83</sup>, ou seja, não há possibilidade de conceber uma divisão da postura intencional em: *original* – relação entre nossos pensamentos com nossos desejos e intenções; e *derivado* – são as representações possíveis a partir da intencionalidade original como (por exemplo) palavras, frases, livros, mapas e programas de computador. Essa difusa postura crítica é contestada pelo autor. Ele está convicto que não é possível estabelecer essa cisão conceitual, alegando que “não há uma maneira baseada em princípios (teoricamente motivada) para distinguir intencionalidade original da intencionalidade derivada” (Dennett, 2009b, p. 343 – nossa tradução).

Outro problema nessa categorização é que não se pode estabelecer um limiar, ainda que teórico, que divide os sistemas intencionais em originais e derivados, visto que há um *continuum* de casos de atribuições e predições legítimas. Por isso, não há intencionalidade de caráter original, mas somente derivada. Apesar de sermos inclinados a conceber que um autor de um artefato intrigante deva possuir intencionalidade intrínseca para ser a fonte da intencionalidade derivada do artefato, segundo o filósofo, não ocorre nada semelhante, pois um “jogador de poker robô que blefa seus criadores parece ser guiado por estados internos que funcionam exatamente como as intenções de um jogador de pôquer humano, e se isso não for intencionalidade original, é difícil dizer por que não” (*Ibidem*).

---

<sup>82</sup> De acordo com o autor, realizar uma interpretação a partir da postura intencional equivale a assumir o “papel de seu guardião, perguntando-nos, de fato: ‘O que eu faria se estivesse na posição deste organismo?’ E aqui nós expomos o antropomorfismo subjacente da postura intencional: tratamos todos os sistemas intencionais como se fossem exatamente como nós – o que, é claro, não são” (Dennett, 1997, p. 29 – nossa tradução).

<sup>83</sup> Os críticos (dentre eles Searle) da teoria dos sistemas intencionais admitem a utilização de tais sistemas sob o jugo da distinção entre duas categorias, a saber: S.I. Original e S.I. Derivado. Segundo essa divisão, somente o primeiro seria uma intencionalidade passível de atribuições e predições genuínas. O segundo, em determinado nível suas predições não são de caráter literal, mas metafórico. (cf. Dennett, 2009b, p. 343).

Voltemos ao exemplo que iniciou este capítulo. Uma bactéria heterotrófica em uma lâmina de vidro ou prato de laboratório. Seus movimentos evitam o contato com um antibiótico em teste em uma extremidade e se projeta em direção a outra extremidade onde se encontram moléculas orgânicas que lhe servem de nutrição. Um observador pode indagar que “ela está procurando (raciocinando) o alimento ou está buscando instintivamente se livrar do extermínio?”. A resposta a essa pergunta é irrelevante, pois, em ambos os casos – e em outros similares – é condicionada a realização de predições e descrição de atribuições. O fato é que, segundo o autor, estamos utilizando para a análise a postura intencional. De modo ainda mais claro, Dennett (1997, p. 25) nos ajuda:

a postura intencional não apenas é uma boa ideia, mas também é a chave para desvelar os mistérios da mente – de todos os tipos de mentes. É um método que explora as similaridades para descobrir diferenças – o grande conjunto de diferenças que se acumularam entre as mentes de nossos ancestrais e as nossas, e também entre as nossas mentes e aquelas de nossos companheiros habitantes do planeta.

No entanto, Dennett (2006) não afirma que sistemas intencionais realmente possuem volições, crença e desejos genuínos, mas apenas que se pode atribuir estados mentais a esses sistemas, melhor dizendo, podemos interpretar, através da observação do comportamento, que determinado ente possui crenças e desejos. Nesse sentido, podemos entendê-lo como uma “estratégia de interpretar o comportamento de uma entidade (pessoa, animal, artefato, etc.), tratá-la como se fosse um agente racional que governasse as suas *escolhas de ação* por uma consideração de suas *crenças e desejos*” (Dennett, 2009b, p. 339 – grifo nosso).

O delineamento do conceito de postura intencional na teoria de Dennett é basilar para as suas pretensões. Ela nos permite realizar descrições de comportamentos que caracterizam e fundamentam a posse de estados mentais. São, segundo o filósofo, indubitavelmente o resultado de um processo evolutivo. Nossos ancestrais, dirá o autor, são robôs, é disso que somos compostos, a intencionalidade que lançamos mão tem origem em uma intencionalidade mais fulcral, proporcionada por um sem número de sistemas intencionais mais rudimentares.

### 3. REAÇÃO AO FUNCIONALISMO HOMUNCULAR

A teoria da consciência de Dennett foi apresentada, até agora, sem maiores críticas. Embora seu argumento pareça, em nosso entendimento, estar no caminho certo, é preciso examinar alguns aspectos nebulosos de sua teoria, apontados por seus críticos – e verificar até que ponto esses supostos equívocos possuem fundamentos. Este é o desígnio maior deste capítulo. Além disso, uma vez estabelecido o contra-argumento do filósofo a essas críticas, nos resta ainda examinar como o autor aborda o problema da identidade pessoal.

De acordo com Dennett (1991a), os termos psicológicos usuais, por serem derivados da psicologia popular, são vagos e imprecisos, mas também são altamente sedutores. De forma bastante eloquente, Teixeira (2008, p. 08) contribui conosco:

A psicologia dificilmente consegue fazer predições exatas do comportamento humano como o faz a mecânica celeste, mas, talvez, apenas algumas aproximações indutivas. Uma possível razão para isso seria o fato de não sabermos se os termos que ela emprega tais como: ansiedade, assertividade, inveja, etc. tenham referente no mundo. Ou seja, ao utilizar essas terminologias na construção de teorias, o psicólogo não sabe muito bem do que está falando: de coisas que existem no mundo ou de ilusões criadas pela sua própria linguagem.

Na teoria dennettiana, a mente se configura como um conjunto de ficções úteis<sup>84</sup> (algo que está no domínio do virtual), que são interpretados por alguns como termos intencionais<sup>85</sup>, mas que na verdade não existem por si só. Como tal, podem ter suporte em uma multiplicidade de organizações materiais, ou seja, não há uma determinação prévia de design. Assim, os conteúdos mentais são, não apenas, construções teóricas, mas também linguísticas (cf. Teixeira, 2008, p. 09). Em termos concretos, isso significa que os estados mentais são formados por uma construção teórica oriunda de termos psicológicos que permitem, enquanto ficções úteis, a realização de uma interpretação de comportamentos complexos – independentemente se são apresentados por humanos ou dispositivos artificiais. Essa construção teórica – os sistemas intencionais – é percebida tanto na natureza quanto no próprio heterofenomenologista.

Dennett (1991a; 1991b) defende que nossas mentes são interpretações do que ocorre em nossos cérebros, e rejeita a concepção de que as ideias possuem existência própria. Diante disso, a teoria dennettiana nega que seja possível reduzir a descrição do comportamento humano – tingidos de termos mentais – a uma cadeia de referentes definidos fisicamente. A descrição das ações humanas não é prescindida de termos intencionais. Neste sentido, cabe asseverar que:

---

<sup>84</sup> Uma espécie de imaginação – no sentido de criação de imagem, ou seja, uma representação – para em seguida responder a esta questão: o que faço com isso?

<sup>85</sup> De acordo com Teixeira (2008), os termos intencionais são termos que extraem sua força no testemunho da experiência subjetiva.

A psicologia parece então enfrentar um grande paradoxo: para se adaptar ao método científico e tornar-se efetivamente uma ciência do psiquismo é preciso que ela exclua de seu escopo a experiência subjetiva e os termos intencionais (com S), ou seja, aquilo que mais acentuadamente constitui sua trama discursiva (Teixeira, 2008, p. 08).

Portanto, a partir do testemunho da minha experiência subjetiva, supostamente, seriam construídos termos (os termos intencionais) que podem ser submetidos a uma extensão para descrever comportamentos. Sua força estaria na experiência subjetiva. No entanto, se quisermos elevar o status da psicologia ao de ciência séria, devemos considerar o abandono da perspectiva em primeira pessoa. Emerge aqui um gravíssimo problema metodológico cuja teoria dennettiana se propõe superar: o que seriam os estados mentais ou *qualia*<sup>86</sup>? As experiências internas (o modo como vemos o mundo) são extremamente sedutoras para estabelecer uma adesão à abordagem em primeira pessoa. Elas parecem insuportavelmente óbvias, mas apresentam facilidade em estabelecer hipóteses difusas como aquela defendida pelos dualistas (cf. Dennett, 1991a, p. 338).

Dennett (1991a) não acredita na existência de dados imediatos, nem enquanto experiências conscientes, nem enquanto *qualia*. Sua tese rejeita o aspecto intuitivo da relação mente/cérebro. Para tanto, se vê diante da necessidade de tecer uma crítica a alguns filósofos contemporâneos que defendem esse ponto de vista. Thomas Nagel é um dos alvos dessa crítica.

Segundo Nagel (2005), o problema da consciência é insolúvel, pois é, como alguns de seus conteúdos específicos, inefável. O filósofo não acredita na redutibilidade de uma abordagem do problema em primeira pessoa para uma de terceira pessoa. Dito de outro modo, não é possível haver uma conciliação entre experiências subjetivas e um discurso científico, objetivo e em terceira pessoa. De acordo com o filósofo, nossa linguagem – sobretudo a linguagem intersubjetiva da ciência contemporânea – não teria competência para alcançar o aspecto fundamental da experiência subjetiva, expresso, por exemplo, na dificuldade que temos de imaginar como é ser um morcego. Ainda que seja possível dominar os aspectos gerais da fisiologia do sistema auditivo desse animal, não poderíamos realizar uma descrição fiel de sua vivência, não poderíamos conhecer de fato suas experiências auditivas<sup>87</sup> – só poderíamos especular sobre o mundo na perspectiva do morcego. No que tange o ser humano, não é tão diferente. Apesar de nossa experiência consciente ser, acredita-se, imediata – sabemos que a temos – não podemos reduzi-la à linguagem. Nesse sentido, a linguagem é insuficiente para

---

<sup>86</sup> Os estados mentais comumente são referenciados como *qualia*, um termo em “latim para qualidades; o singular é *quale*, geralmente pronunciado kwah-lay” (Dennett, 1991a, p. 338 – nossa tradução). Embora, haja uma discussão sobre se realmente todo estado mental seja dotado de *qualia*.

<sup>87</sup> De acordo com a biologia, os morcegos têm a capacidade de realizar ecolocalização, ou seja, executa discriminações perceptivas de natureza auditiva tão acuradas quanto aquelas que nós realizamos com a visão.

descrevê-la em sua plenitude – já que cada ser humano, aparentemente, realiza uma representação única do mundo (cf. Nagel, 2005, p. 246).

Segundo Teixeira (2008), é crucial, na teoria da consciência dennettiana, demonstrar que os *qualia* não podem ser apreendidos intuitivamente. Essa é uma hipótese difícil de se defender, haja vista que grande parte dos teóricos da mente os consideram como estados fenomenais intrínsecos ao sentimento de um *eu*. Eles entendem que afirmar que os *qualia* não existem seria análogo a negar que realizamos vivências internas. No entanto, Dennett não defende a inexistência dos *qualia*, como muitos erroneamente o interpretam – sua abordagem não é ontológica, mas epistemológica. Em termos concretos, isso significa que considera válido o sentimento de um *eu* e suas experiências, mas nega que sejam inefáveis. Assim, podemos concluir que “os sentimentos de si só passam a existir na medida em que são reportados pelo sujeito que os tem” (cf. Teixeira, 2008, p. 24). Sem embargo, seria possível descrevê-los na medida em que são examinados numa abordagem em terceira pessoa, ou seja, heterofenomenológica. A fim de facilitar essa interpretação, Teixeira (2008, p. 11) traça uma importante linha tangencial entre a abordagem dennettiana e o computacionalismo de Turing, afinando nosso entendimento:

O critério proposto por Turing para saber se algo ou alguém pensa é baseado na possibilidade de atribuição de estados mentais a esse algo ou alguém. Seu critério é operacional, ele evita a necessidade de discutir a natureza dos estados mentais que seriam responsáveis por quaisquer peculiaridades de comportamento. A premissa fundamental de Turing [...] é que se o comportamento de uma máquina for indistinguível daquele de um ser humano, não há nenhuma razão que nos impeça de atribuir a essa máquina estados mentais.

Apesar de ser um caso específico de aplicação desse critério, o teste de Turing é de suma importância, porquanto, se configura na aplicação de casos particulares do comportamento linguístico. O pensamento dennettiano bebe dessa fonte, se caracterizando também por ser de cunho pragmático e operacional, ou seja, devemos conceber que na aplicação do critério operacional se estabelece que “estados mentais passam a ser vistos como termos teóricos que, uma vez atribuídos a segmentos de comportamentos, tornam estes inteligíveis para nós” (*Ibidem*). Neste ínterim, cabe asseverar que:

O sistema intencional – e com ele os elementos que o compõem, todos os derivados da psicologia popular – é uma medida de nossa ignorância para prever o comportamento de um organismo ou dispositivo complexo. A complexidade é aqui a chave de tudo: se um animal ou robô se comporta de maneira tão complexa que para podermos montar uma explicação ou uma história de seu comportamento, precisamos lançar mão de intenções, crenças, desejos, etc., é perfeitamente legítimo, no entender de Dennett, atribuir a essa criatura uma vida mental (Teixeira, 2008, p. 13).

Nessa perspectiva, a psicologia dennettiana é assimilada como uma estratégia de predição, fornece uma supremacia evolutiva aos organismos que se caracterizam por esse

atributo, ou seja, o fato pelo qual os seres humanos possuem a capacidade de lembrar, acreditar, conhecer, eles são agraciados com grande poder de predição, tanto em relação ao comportamento de outros humanos, quanto ao de outros animais (cf. Teixeira, 2008, p. 13). Assim, podemos conceber que, em um nível mais elevado, “a inteligência é a capacidade de um organismo ou máquina de prever a previsão do outro e basear seus comportamentos futuros nessa previsão, como uma espécie de resposta” (*Ibidem*).

Não há um administrador central no cérebro ou um lugar especial onde se encontra a mente. A ênfase dada à intuição, na edificação de teorias, causa uma patologia teórica – leva os filósofos a ignorar que suas hipóteses ainda consideram, direta ou indiretamente, a existência de que em algum lugar, convenientemente oculto no cérebro, se encontra o já comentado teatro cartesiano. Vimos como essa não pode ser uma boa ideia, mas ainda há pensadores fascinados por ela (cf. Dennett, 1991a, p. 39). Chalmers ao defender a irreducibilidade do caráter consciente da experiência, demonstra flertar com essa perspectiva.

### **3.1 O embate com Chalmers: será que há o problema difícil?**

O embate filosófico entre o dualismo naturalista de David Chalmers e o funcionalismo homuncular de Dan Dennett é claramente uma divergência de princípios<sup>88</sup>. O conflito entre seus argumentos reside no fato de que o estadunidense não acredita que tomar a intuição da consciência como um dado seja um bom caminho para tentar explicá-la, enquanto que o australiano defende que a partir da aceitação de uma intuição podemos construir uma teoria, ou seja, só podemos examinar a consciência pressupondo a sua existência.

Consoante Chalmers (1996), a consciência é um truísmo, um dado intuitivo sem possibilidade de refutação. Sua força ontológica reside justamente em seu caráter intuitivo. Nessa perspectiva, o autor a descreve como um problema difícil (*hard problem*), em comparação com o problema fácil – aquele que se reflete nos trabalhos da neurociência e da ciência cognitiva, por exemplo. Além disso, o filósofo australiano estabelece a distinção<sup>89</sup> entre mente psicológica e mente fenomenal.

Ao tomar a mente como explicação ou causa de determinado comportamento de uma entidade, estamos diante da concepção de mente psicológica chalmersiana. Nesse sentido, um evento mental psicológico é a atividade de memorização de um processo cognitivo, uma espécie de instanciação de processamento de “dados” (como por exemplo, representações mentais de

---

<sup>88</sup> De acordo com Leal-Toledo (2006), do ponto de vista da análise da mente, ou seja, aquela que exclui a consciência, Chalmers se comporta como um funcionalista.

<sup>89</sup> Essa distinção nos ajuda compreender a relação de Chalmers com a psicologia cognitiva e o funcionalismo, por um lado, e com a fenomenologia da consciência, por outro lado.

objetos e as operações realizadas com essas representações). Tais processos não necessariamente são acompanhados pela mente fenomenal, ou seja, a experiência consciente – uma caracterização da mente pelo modo que sentimos nossos processos mentais. A noção psicológica de mente, contrariamente, pode ser delineada a partir da identificação do que a mente é capaz de fazer. Esses dois aspectos são irrefutáveis, e uma teoria séria da consciência, afirma Chalmers, não deve se abster de nenhum deles. Assim, há uma dualidade inerente a processamentos mentais conscientes (a noção de percepção, por exemplo, é passível de ser assimilada psicologicamente, seja como uma experiência introspectiva ou uma imagem mental de um objeto, uma intuição, uma sensação exclusiva, do objeto representado. Os eventos mentais fenomenais, apesar de não serem redutíveis, estão intrinsecamente relacionados à instância psicológica – à ocorrência de um processo cognitivo (cf. Chalmers, 1996, p. 24-31).

Em seu entendimento, essa distinção é fundamental para a análise da ontologia da inteligência consciente, pois, segundo o filósofo, é comum estabelecer uma confusão entre esses campos de investigação, de tal modo que “alguém pode começar questionando o problema da experiência qualitativa, mas acaba dando uma resposta sobre como um organismo consegue relatar seus estados internos” (Leal-Toledo, 2006, p. 124).

O nível de complexidade de descrição da natureza que constitui a experiência consciente é designado por Chalmers por *problema difícil*. Esse óbice, segundo o autor, se eleva ao nível de exaurir o conjunto de recursos disponíveis às análises funcionalistas, ou seja, “o que torna complexo o problema difícil e quase único é que ele vai além dos problemas sobre o desempenho das funções” (Chalmers, 1996, p. 12 – nossa tradução). Sua crítica ao funcionalismo dennettiano é caracterizada pela acusação de que o estadunidense depositou sua indústria na abordagem do problema fácil, deixando em aberto o problema difícil.

Ao tentar explicar a natureza da inteligência consciente, nos deparamos com a hipótese de que ao experienciar o mundo, concomitantemente, articulamos um processamento de informação, composto tanto de aspectos físicos quanto subjetivos. O problema difícil, por perpassar as explicações da consciência, por transpor a explicação das estruturas funcionais que parecem permear a experiência consciente (categorização, discriminação, volições, etc.), deixa em aberto uma questão desconcertante: como é possível a performance dessas funções ser acompanhada por experiências? Desse modo, a questão basilar para entender os problemas que envolvem a consciência é explicar como um ser ou um artefato consegue realizar experiências mentais – sabemos muito pouco sobre tais experiências.

De acordo com Chalmers (1996), as lacunas não foram preenchidas. Há uma divisão subjacente que estabelece uma distinção fundamental para a análise do problema: há o aspecto

objetivo – que está relacionado às funções; e o aspecto subjetivo – que está relacionado à experiência mental. Se faz necessário, portanto, a edificação de uma hipótese que englobe a dificuldade de interação entre esses aspectos. Para tanto, de acordo com o filósofo, é preciso estabelecer novas abordagens – os métodos reducionistas dariam conta dos aspectos funcionais, mas são ineficientes para explicar as questões sobre as experiências introspectivas. Vejamos como Dennett refuta esse argumento.

Consoante Dennett (1991a), não há outro modo de abordar o problema da consciência, se não com o fisicalismo. Segundo o filósofo, o dualismo foi superado pelos progressos dos trabalhos da neurociência e da inteligência artificial. O cérebro é uma estrutura que trabalha com processamentos paralelos e que, mais importante, não culmina em um fluxo contínuo de consciência (o teatro cartesiano). Os dualistas admitem tal fluxo. Podemos resumir a crítica dennettiana de como os dualistas concebem a mente e a consciência, numa reformulação, talvez grotesca, da parábola da caverna de Platão: as sombras seriam a consciência, os prisioneiros o dualista, os transeuntes que conversam e levam estatuetas seriam a atividade cerebral e o fugitivo, diria Dennett, seria o heterofenomenologista – que percebe que as vivências dos cárceres é condicionada pela estrutura da caverna.

A tese funcionalista homuncular descarta a possibilidade de haver uma reunião entre os vários circuitos especialistas. O que existe na verdade é o trabalho individual e paralelo de cada circuito neural que procura desenvolver o seu próprio rascunho narrativo, que não necessariamente fará parte do texto publicado ou será refletido em um comportamento. Qual rascunho comandará o comportamento? Isso vai depender dos *inputs* ao que o indivíduo foi submetido (cf. Leal-Toledo, 2006 p. 127). Neste íterim, em relação às contradições do ponto de vista do teatro cartesiano, cabe asseverar que:

O cérebro então deveria fazer sempre duas coisas: primeiro processaria a informação de cada coisa separadamente em suas respectivas áreas cerebrais e depois reuniria tudo para formar uma experiência consciente do conjunto. É o chamado problema da ligação (*the binding problem*). Só que não há no cérebro um lugar onde tudo se reúne ou onde tudo deve passar e, para Dennett, nem há necessidade de haver. Também não há necessidade alguma de toda a informação que já foi processada em áreas separadas do cérebro seja mais uma vez processada para se tornar uma experiência consciente, como se tudo tivesse que ser feito duas vezes (Leal-Toledo, 2006, p. 126).

Descrever a visão dennettiana sobre o problema da consciência é extremamente difícil. Talvez uma análise de uma situação corriqueira nos ajude a torná-la um pouco mais intuitiva: podemos nos lembrar daquelas situações em que só percebemos um som quando ele cessa – antes disso não estávamos conscientes dele e, portanto, não o ouvíamos. Essa situação só é inusitada para os partidários do modelo de teatro cartesiano. De acordo com o modelo de rascunho múltiplos, já estava sendo elaborada uma narrativa durante a execução do barulho,

mas ainda não exercia nenhuma influência no julgamento e em nossa memória sobre os fatos que nos circundam naquele momento. A narrativa só será concluída quando o barulho cessar, é quando se torna consciente. É nesse sentido que podemos afirmar que “não há fatos fixos a respeito do fluxo de consciência que sejam independentes das sondagens particulares” (Dennett, 1991a, p. 138 – nossa tradução).

Essa postura é indubitavelmente um modelo inovador de tratar os *qualia*, mas não implica em uma negação radical de sua existência. O que é negado efetivamente é que eles existem como propriedades qualitativas intrínsecas. O que realmente existe são “nossos julgamentos, nossas decisões, nossa memória, nossos pensamentos sobre os *qualia* e não os *qualia* enquanto tais” (Leal-Toledo, 2006, p. 127 – grifo nosso). Quando fazemos uma descrição do tipo “ontem a maré estava agitada” não estamos realizando uma publicação de uma propriedade de nossa mente, mas, contrariamente, estamos rememorando um julgamento de que “ontem a maré estava agitada”. De forma bastante eloquente, Dennett (1991a, p. 364 – nossa tradução) afirma:

Você parece achar que há uma diferença entre o pensamento (julgamento, decisão, defesa convicta) de que alguma coisa é rosa para você, e o fato de que algo realmente parece ser rosa para você. Mas não há diferença. Algo como “realmente parecer” não pode ser considerado um fenômeno para além do fenômeno de julgar de uma maneira ou de outra que algo é como é.

Chalmers (1996) não vê com bons olhos a postura dennettiana, pois entende que o estadunidense evitou o problema difícil, e se limitou apenas a explicar como se estabelece nossa capacidade de reportagem de estados internos e de direcionamento. Em resposta, Dennett afirma que sua teoria é a melhor maneira de explicar a consciência, pois “apenas uma teoria que explicasse os eventos conscientes em termos de eventos inconscientes poderia explicar a consciência” (Dennett, 1991a, p. 454 – nossa tradução). Além disso, acusa Chalmers de circularidade, ou seja, de delinear uma certa noção de consciência e em seguida procura demonstrar a veracidade dessa concepção (cf. Leal-Toledo, 2006, p. 128). Isso fica evidenciado quando o australiano adverte que uma verdadeira teoria da consciência deve levá-la a sério. Portanto, tratar a consciência com seriedade é tratá-la da sua perspectiva. Esse é o cerne da divergência com o pensamento dennettiano, pois é óbvio que o funcionalismo homuncular não parte desse pressuposto.

Por conseguinte, Chalmers (1996) desenvolve uma noção de consciência que culmina necessariamente no problema difícil, ou mais precisamente em seu dualismo naturalista da propriedade. É nesse sentido que o filósofo salienta quando diz que “o verdadeiro argumento do livro é que se alguém leva a consciência a sério, deveria chegar à posição que exponho”

(Chalmers, 1996, p. 13 – nossa tradução). Não surpreende a acusação de circularidade ao dualismo naturalista, haja vista que culmina na “conclusão de que a consciência é uma entidade básica e primitiva do nosso mundo, que não poderia ser explicada em termos de outras entidades” (cf. Leal-Toledo, 2006, p. 128). Sem embargos, Chalmers não explica o que é a consciência. Para o autor, intuímos a consciência indubitavelmente, e o máximo que podemos fazer é aceitar a sua existência, ou seja, estamos diante de uma certeza ontológica calcada em um ceticismo epistemológico.

O caminho de Dennett (1991a) é diferente. Descreve uma teoria que explica a consciência por meio de processos inconscientes. Isso implica uma noção de zumbis que destoa daquela desenvolvida por Chalmers. Por ora, vejamos o que Dennett (1995c, p. 322-23 – nossa tradução) diz a esse respeito:

Algumas vezes, os filósofos agarram-se desesperadamente a hipóteses insuportáveis, correm até um penhasco, e atiram-se de cabeça. Então, como personagens de desenho animado, ficam parados no ar, até que percebem o que fizeram e a gravidade começar a funcionar. Isso aconteceu com os filósofos que adotaram o conceito de zumbi, uma noção estranhamente atraente que soma, em uma pilha só, quase tudo que eu acho que está errado no pensamento atual a respeito da consciência.

Com efeito, Dennett (1999) acredita que canalizar esforços para discussões sobre zumbis é desnecessário, para não dizer ridícula. Ela sequer deveria ser cogitada, pois entende que é um conceito inútil e que flerta com antigos preconceitos. Contudo, a intuição – *hunch*<sup>90</sup> – dos zumbis é altamente sedutora (o próprio Dennett afirma não ser imune ao canto dessa sereia), mas considera que bons argumentos superam intuições: intuitivamente a terra nos parece plana, mas construímos bons argumentos para demonstrar que ela participa de uma realidade bastante diferente; intuitivamente nos parece que não há distinção entre um objeto em repouso absoluto e um objeto em movimento retilíneo uniforme, mas bons argumentos demonstraram que essa postura é equivocada. De modo similar, para muitos “ainda parece que esta teoria mecanicista da consciência deixa algo de fora, mas isso, obviamente, é uma ilusão” (Dennett, 1999, p. 09 – nossa tradução). Ainda mais grave, nenhum crítico, até agora, explicou o que é esse “algo”.

O novo que surge com Chalmers (1996) é a hipótese da impossibilidade de engendrar uma explicação reducionista para a natureza da inteligência consciente. O australiano nos apresenta um dualismo naturalista acreditando que a experiência consciente deve ser uma entidade que participa daquele quadro formado pelas características fundamentais do universo – como por exemplo, a energia, a massa e o espaço-tempo. Entrementes, não apresenta uma relação necessária com sua base física, ou seja, seria um fator suplementar que não deriva de

---

<sup>90</sup> O termo *hunch* é de difícil tradução, pode significar um palpite ou uma crença intuitiva em algo.

sua estrutura física. O dinheiro físico e as obras de arte são exemplos que o autor recorre para realizar uma analogia. O valor da cédula, por exemplo, não supervêm necessariamente de sua composição física, haja vista que há cédulas falsificadas. Um quadro como *o grito* de Edvard Munch (1863-1944), se for reproduzido, não terá o mesmo valor estético – seu valor deriva de algo que não é o quadro em si.

Como vimos, Chalmers (1996) nos pede para considerar seres que, ainda que possuam todas as características de um ser consciente, não seriam conscientes. Essa hipótese é levantada em função do problema difícil, na medida em que se admite que “a experiência consciente legítima é autorreflexiva e que a autorreflexão não deriva de dispositivos materiais suplementares que funcionariam como leitores dos relatos de nossos próprios desempenhos” (Teixeira, 2008, p. 26). Segundo essa perspectiva, a superveniência lógica dos estados qualitativos é a raiz do fisicalismo em geral, bem como do reducionismo.

Consoante Chalmers (1996), os estados mentais seriam um tipo de fenômeno único da natureza, e, portanto, descartam a possibilidade de sua reprodução artificial. Esse entendimento o conduz a edificar um dualismo naturalista. A busca pela refutação ao fisicalismo leva o australiano a formular o argumento zumbi – a suposição de uma entidade que apresentaria tanto os aspectos físicos quanto comportamentais de um indivíduo humano. Não obstante, como lhe faltaria a consciência de sua atividade, não poderia ser considerado um ser humano. A simples possibilidade, lógica ou metafísica, de existir zumbis permite a exposição dos limites do fisicalismo. Ironicamente, de acordo o filósofo, a mesma noção de superveniência lógica é capaz de servir de base para estabelecer uma refutação as concepções fisicalistas: se for possível imaginar um mundo que seja fisicamente igual ao nosso, mas em que não exista a possibilidade da consciência supervir, implicaria que ela não é uma necessidade física, ou seja, a consciência não é uma superveniência lógica (cf. Leal-Toledo, 2006, p. 125). Todavia, se tal mundo for contraditório, de algum modo, sua hipótese estará refutada. Chalmers (1996, p. 94-95 – nossa tradução) nos traz um panorama que facilita mais o entendimento de sua hipótese:

A maneira mais óbvia (embora não a única) de investigar a superveniência lógica da consciência é considerar a possibilidade lógica de um zumbi: alguém ou algo fisicamente idêntico a mim (ou a qualquer outro ser consciente), mas totalmente desprovido de experiências conscientes. No nível global, podemos considerar a possibilidade lógica de um mundo zumbi: um mundo fisicamente idêntico ao nosso, mas no qual não há experiências conscientes [...] então vamos considerar meu gêmeo zumbi. Essa criatura é molécula por molécula idêntica a mim, e idêntica em todas as propriedades de baixo nível postuladas por uma física completa, mas ela carece inteiramente de experiência consciente [...] sensações de ver os verdes das árvores do lado de fora, ter experiências de sabor agradáveis ao mastigar uma barra de chocolate e sentir uma sensação de dor no meu ombro direito.

E continua:

O que está acontecendo no meu gêmeo zumbi? Ele é fisicamente idêntico a mim, e podemos supor que ele está inserido em um ambiente idêntico. Ele certamente será idêntico a mim funcionalmente: ele estará processando o mesmo tipo de informação, reagindo de maneira semelhante às entradas, com suas configurações internas sendo modificadas adequadamente e resultando em um comportamento distinguível. Ele será psicologicamente idêntico a mim [...] Ele estará percebendo as árvores lá fora, no sentido funcional, e provando o chocolate, no sentido psicológico. Tudo isso decorre logicamente do fato dele ser fisicamente idêntico a mim, em virtude das análises funcionais das noções psicológicas. Ele estará até mesmo "consciente" nos sentidos funcionais descritos anteriormente — estará acordado, capaz de relatar o conteúdo de seus estados internos, capaz de focalizar a atenção em vários lugares e assim por diante. Só que nada desse funcionamento será acompanhado por qualquer experiência consciente real. Não haverá sensação fenomenal. Não há nada como ser um zumbi (*Ibidem*).

Embora a proposta de Chalmers (1996) pareça estar seguindo os parâmetros de um experimento de pensamento destrutivo, que objetiva refutar as teorias fisicalistas, na verdade está mais próximo daquilo que Dennett denomina de *spear crutches*, uma muleta de lança. O australiano sugere que devemos verificar se a hipótese zumbi tem fundamento e só após a constatação dessa possibilidade lógica, edificar uma teoria que explique a natureza da inteligência consciente.

De acordo com Teixeira (2008), os zumbis são frutos de uma ontologia estranha e não apresentam uma vida consciente. Esse experimento mental tem a finalidade de fundamentar o problema difícil, bem como refutar o fisicalismo<sup>91</sup>. O dualismo de Chalmers se distancia do cartesianismo por este último não considerar a possibilidade de que os estados mentais sejam supervenientes de autômatos. Embora o australiano não negue categoricamente essa possibilidade, ele critica a hipótese de que a analogia funcional seja suficiente para explicar os estados conscientes de cunho metafísico. Nesse sentido, se mantivermos o primado dos dados intuitivos para engendrar uma teoria da consciência, há poucos motivos para negar que uma máquina, portadora de uma inteligência artificial e que demonstre comportamento humano, vivencie experiências conscientes. Há outras questões que emergem a partir dessa constatação: o dualista poderia afirmar que a máquina em questão seria no máximo um zumbi. Mas um zumbi é possível, ainda que metafisicamente? Uma entidade que apresenta toda uma anatomia de um ser humano, bem como todo comportamento possível de sua natureza (agem, conversam, etc.) e que, portanto, poderia passar com idoneidade no teste de Turing. Esse é o zumbi que Chalmers nos pede para considerar. O único agravante que os distingue de um ser humano é que um zumbi não teria experiências conscientes. Mas como isso seria possível? Se um zumbi apresenta um comportamento que seria indistinguível de um ser humano, o que nos dá o direito de não lhe atribuir estados mentais?

---

<sup>91</sup> Nesse sentido, aparentemente, está mais próximo da noção de experimento de pensamento platônico de Brown.

Para facilitar nossa compreensão do problema com zumbi de Chalmers, considere a seguinte situação: um determinado indivíduo se depara com um zumbi, que está sendo submetido ao teste de Turing, e tece a seguinte questão: “você tem experiências conscientes?” ou “você tem consciência daquilo que acabaram de te perguntar?” Sim ou não, seriam as possibilidades de respostas. A resposta seria irrelevante, haja vista que nos dois casos possíveis de respostas, já está implícito o estado consciente do suposto zumbi (esteja ele mentindo ou não). Se não fosse capaz de fornecer nenhuma resposta, o zumbi não passaria no teste de Turing, mas também não faria tudo o que um ser humano é capaz de fazer (com exceção de vivenciar estados conscientes como nos sugere Chalmers), o que o tornaria uma impossibilidade lógica.

Consoante Dennett (1999), os zumbis são logicamente inconcebíveis. A teoria da superveniência dos estados conscientes é herdeira direta dos mesmos problemas do cartesianismo. De forma bastante eloquente, Teixeira (2008, p. 26) contribui conosco:

Descartes sustentava que a duplicação de características materiais e funcionais de um ser humano poderia ser condição necessária, mas não suficiente, para se replicar a vida mental humana. Um autômato bem construído poderia vir a fazer tudo o que um ser humano faz, mas nunca se igualaria a ele: seria, no máximo, uma proeza de engenharia, algo que, contudo, não teria alma. E não poderíamos substituir esta palavra por "experiência consciente"? Neste sentido, o autômato de Descartes não é muito diferente do zumbi de Chalmers.

Considere atribuir a um organismo dois estados mentais acompanhados pelos seus respectivos comportamentos. Explicar a estrutura física do organismo em questão não é suficiente para estabelecer essas atribuições – é necessária uma observação de suas características globais. No entanto, somos inclinados ao erro de considerar as características globais como de caráter adicionais. Não faz sentido (ainda que metafisicamente) acreditar que é possível extrair a saúde, por exemplo, e manter inalterado o funcionamento de todos os órgãos. Também não nos parece lógico a remoção ou lesão de um desses órgãos e esperar a preservação da saúde do organismo. Por analogia, a consciência também não parece poder violar essa regra lógica.

Dennett (1991a), eleva o tom de sua crítica ao palpite zumbi. Será que uma das pessoas que você conhece é um zumbi (no sentido “técnico” dos filósofos)? O termo *zumbi* é oriundo, segundo o autor, do folclore do vodu haitiano<sup>92</sup>. Se aproxima daquele conceito explorado no cinema<sup>93</sup>. Mas Chalmers nos convida a imaginar uma categoria diferente de zumbi. Como tudo

---

<sup>92</sup> Nesse contexto, é um ritual de punição para uma pessoa que cometeu “algum delito e foi condenada a se arrastar, resmungando e olhando com olhos mortos, fazendo negligentemente as ordens de algum sacerdote vodu ou xamã” (Dennett, 1991a, p. 270 – nossa tradução).

<sup>93</sup> O antropólogo Wade Davis afirma “que havia decifrado o mistério dos zumbis vodu, e em seu livro *The Serpent and the Rainbow* (1985) descreveu a poção neurofarmacológica preparada por praticantes de vodu que pode colocar seres em um estado de morte; depois de serem enterrados vivos por vários dias, essas pessoas infelizes às

o que podemos perceber das pessoas que nos são próximas são os seus comportamentos, infere-se daí que algumas delas poderiam ser zumbis. Isso já seria um bom motivo para o estadunidense ser neutro em relação a esse conceito.

Sem embargo, apesar da teoria dennettiana não realizar suposições ontológicas da consciência, ela canaliza sua atenção aos seres humanos adultos normais, haja vista que se a inteligência consciente envolve alguma entidade, esse seria o possível hospedeiro. Só após essa análise podemos ampliar e “voltar nossa atenção para a consciência (se houver) de outras espécies” (Dennett, 1991a, p. 270 – nossa tradução). O autor evita traçar uma linha tangencial entre o conceito zumbi e sua teoria, afinando nosso interesse por ora:

Shakey era um zumbi particularmente rude, mas agora podemos imaginar um zumbi mais realista e complexo, que monitora suas próprias atividades, incluindo até suas próprias atividades internas, em uma espiral ascendente indefinida de reflexividade. Chamarei essa entidade reflexiva de zimbo. Um zimbo é um zumbi que, como resultado do automonitoramento, tem estados informativos de ordem superior internos (mas inconscientes) que se referem a seus outros estados informativos de ordem inferior. (Não faz diferença para este experimento mental se um zimbo é considerado um robô ou uma entidade humana – ou marciana.) Aqueles que acreditam que o conceito de um zumbi é coerente certamente devem aceitar a possibilidade de um zimbo. Um zimbo é apenas um zumbi de complexidade comportamental, graças a um sistema de controle que permite a autorrepresentação recursiva (Dennett, 1991a, p. 310 – nossa tradução).

Ao invés de considerar o zumbi dos filósofos o autor oferece “esta parábola dos zimbões irônico, uma vez que não acho que o conceito de um zumbi ou as categorias psicológicas populares de pensamentos de ordem superior podem sobreviver, exceto como relíquias de um credo ultrapassado” (Dennett, 1991a, p. 313-14 – nossa tradução). Nesse sentido, o filósofo declara que é impossível diferenciar um zumbi de um ser humano que apresenta estados conscientes, por implicar que a consciência seria – em um sentido claramente ridículo – epifenomenal. Destarte, em vez disso, o epifenômeno no sentido de Huxley<sup>94</sup> é mais conveniente. Não obstante, hajam maneiras de estabelecermos diferenças cruciais entre seres humanos reais e zumbis (quem sabe, por exemplo, seus cérebros sejam de uma tonalidade amarelada), elas não parecem estar calcadas em diferenças funcionais (cf. Dennett, 1991a, p.

---

vezes são exumadas e recebem um alucinógeno que causa alguma desorientação e amnésia. Como resultado do alucinógeno ou do dano cerebral causado pela privação de oxigênio durante o sepultamento, eles de fato se arrastam mais ou menos como os zumbis nos filmes e, às vezes, podem ter sido escravizados. Por causa da natureza sensorial das alegações de Davis (e do filme vagamente baseado em seu livro novelístico), suas descobertas encontraram uma corrente de ceticismo em alguns setores, mas elas são bem refutadas em um segundo livro, mais acadêmico, *Passage of Darkness: a etnobiologia do zumbi haitiano (1988)*” (*Ibidem*).

<sup>94</sup> O termo *epifenômeno* aparece pela primeira vez vinculado a uma patologia, em 1706, remete a ideia de uma aparência secundária ou sintoma” de uma doença. É nesse sentido que o biólogo evolucionista Thomas Huxley (1874), apesar de ter sido possivelmente o primeiro a utilizar o termo no uso atual, numa aplicação relativamente clara e sem problemas, mas irrelevante para os argumentos filosóficos (cf. Dennett, 1991a, p. 401-02).

405). A respeito da possibilidade de existir zumbis, Dennett (1991a, p. 406 – nossa tradução) nos traz um panorama que facilita mais o entendimento de seu pensamento:

Há outra maneira de abordar a possibilidade de zumbis e, em alguns aspectos, acho que é mais satisfatória. Os zumbis são possíveis? Eles não são apenas possíveis, eles são reais. Somos todos zumbis. Ninguém está consciente – não da maneira sistematicamente misteriosa que apoia doutrinas como o epifenomenalismo! Não posso provar que não existe tal tipo de consciência. Eu também não posso provar que gremlins não existem. O melhor que posso fazer é mostrar que não há motivação respeitável para acreditar nisso.

A intuição zumbi deve ser descartada, pois não há o substrato que surgiria ao compararmos um ente análogo a um ser humano, mas que estaria privado de experiências conscientes, e um ser humano real. Não há diferença entre seres humanos e zumbis. Entrementes, o filósofo nos fala de um tipo de zumbi que, a princípio, é indistinguível de um ser humano. Nessa perspectiva, percebemos que, de certo modo, há uma aceitação do argumento zumbi. Então a consciência é um epifenômeno, já que zumbis físicos são possíveis? Surpreendentemente estamos diante de um dos maiores críticos do argumento zumbi confirmando a sua possibilidade. Como vimos acima, o estadunidense afirma ser flexível à intuição zumbi, o que está sendo corroborado agora. Mas considerar a hipótese zumbi não implica aceitar o epifenomenalismo ridículo, haja vista que nesse tipo a consciência não exerceria um papel causal e é descartada a hipótese de seu surgimento ser derivado da evolução biológica. É nesse sentido que os zumbis não devem ser levados a sério. Com efeito, há uma maior canalização da crítica dennettiana para o epifenomenalismo em detrimento da concepção de zumbi a ele inerente.

### 3.2 “Cerceando<sup>95</sup>” Searle

John Searle (1998; 2006) tece uma dura crítica ao fisicalismo, ao delinear uma teoria calcada na irreduzibilidade ontológica da consciência, cuja essência é indiscutivelmente subjetiva, mas que nem por isso deixaria de estar em consonância com a ciência padrão. Nesse sentido, segundo o filósofo, o fisicalismo erra – sendo o funcionalismo homuncular dennettiano o exemplo maior desse erro – por não considerar os aspectos subjetivos da consciência. Ao defender uma ontologia da consciência, o pensamento searleano, estranhamente, estabelece uma privação de uma possível redutibilidade em relação ao cérebro, entendendo que a consciência é biologicamente natural. Assim, se por um lado, o dualismo, ao realizar uma

---

<sup>95</sup> Esse termo é usado com o intuito de enfatizar o tom ríspido que caracteriza o debate entre Dennett e Searle. Algo similar ao que Dennett, em *Brainstorms*, realiza ao usar o verbo “esfolar” (to skim) para tecer uma crítica ao behaviorismo skinneriano na secção Skinner “esfolado” (Skinner Skinned), um trocadilho de difícil tradução (cf. Dennett, 2006 p. 95).

diferenciação ontológica entre mente e cérebro, não resolve o problema da coisa-mente – por não conseguir estabelecer uma descrição inteligível dessa interação. Por outro, o fisicalismo comete o estranho erro de reduzir a mente a alguma outra instância ou simplesmente aniquilá-la – o que, segundo o filósofo, é a proeza de Dennett (cf. Searle, 1998, p. 25).

O que leva Dennett e os funcionalistas a esse tipo de erro, segundo Searle (2006), é a tentativa de naturalizar os estados mentais a partir de uma redução ontológica a processos físicos. A origem desse equívoco, segundo o autor, reside no fato de que no fisicalismo em geral – sobretudo no funcionalismo homuncular dennettiano – há a inclinação em macular o aspecto ontológico do mental, expresso em uma negação de sua existência ou reduzindo sua relevância. Searle (1995, p. 364-65), continua em sua análise acerca de um panorama depreciativo e negativo do mentalismo argumentando sobre este que:

[...] ninguém jamais chegou a essas posições após um exame atento dos fenômenos em questão. Ninguém jamais considerou sua própria e terrível dor ou sua mais profunda preocupação e concluiu que não passavam de estados de uma máquina de Turing, ou que poderiam ser inteiramente definidos em termos de suas causas e efeitos, ou que atribuir tais estados a si mesmo era apenas uma questão de assumir uma certa atitude para consigo mesmo. Em segundo lugar, ninguém pensaria em tratar desse modo outros fenômenos biológicos. Se alguém estivesse fazendo um estudo das mãos, dos rins, ou do coração, simplesmente presumiria a existência das entidades em questão, para então prosseguir o estudo de sua estrutura e função.

Com efeito, Searle (2006) enumera as hipóteses fisicalistas que, no seu entendimento, são consideradas inadmissíveis. De sorte que o funcionalismo dennettiano está nas fileiras da crítica searleana, pois, apesar de não negar os termos mentalistas, entende que não devemos considerar sua existência ontológica, haja vista que são “simplesmente um vocabulário útil para explicar e prognosticar o comportamento, mas não para ser tomado literalmente, como se remetesse a fenômenos psicológicos reais, intrínsecos, subjetivos” (Searle, 2006, p. 16).

Todavia, Dennett (1991a; 1999) tece uma contraofensiva à crítica searleana. Não só canaliza um ataque a um dos pilares do naturalismo biológico, ou seja, ao caráter ontológico da subjetividade, mas também deixa claro que esse ponto é a maior divergência entre sua concepção de consciência e a de Searle. Na passagem a seguir podemos perceber o teor desse embate:

John Searle e eu temos uma profunda divergência sobre como estudar a mente. Para Searle, é tudo realmente muito simples. Existem essas intuições fundamentais, testadas pelo tempo, que temos sobre a consciência e qualquer teoria que as desafie é simplesmente absurda. Eu, por outro lado, penso que o problema persistente da consciência vai permanecer um mistério até que encontraremos algumas dessas óbvias intuições mortas e mostremos que, a despeito das primeiras impressões, ela é falsa! Um de nós está redondamente enganado e os riscos são altos. Searle vê minha posição como ‘uma forma de patologia intelectual’; ninguém deveria surpreender-se ao saber que o sentimento é mútuo. Searle tem a tradição ao seu lado. Como ele diz, minha visão é notavelmente contra-intuitiva no início. Mas, a visão dele também tem alguns

problemas que emergem somente após uma análise um tanto quanto minuciosa. Agora, como procedemos? Cada um de nós procura construir argumentos para demonstrar nosso ponto de vista e comprovar que o outro lado está errado (Dennett *apud* Searle, 1998, p. 133-34).

Dennett (1991a) é categórico. Não há uma estrutura mental tal qual é concebida pela psicologia popular. A simples admissão dessa ideia implicaria em uma redução ontológica da consciência. No funcionalismo homuncular estamos diante de uma interessante teoria epistemológica da consciência.

O pensamento searleano nos apresenta a tese da causação cujo resíduo nos leva a crer que há uma redução dos *qualia* aos processos cerebrais, ou seja, a fenomenologia pura seria descrita através de uma análise dos processos cerebrais e não teria poderes causais além daqueles observados nesses processos. Nessa perspectiva, as capacidades causais da consciência se limitam ao substrato neuronal. Essa hipótese apresenta problemas de extrema complexidade, haja vista que há a possibilidade de interpretá-la como uma defesa da identidade das capacidades causais da consciência e do cérebro, de tal modo que podemos rotulá-la de reducionismo – coisa que é veementemente descartada por Searle (cf. Prata; Lima Filho, 2013, p. 196). Como há um estreito vínculo entre cérebro e consciência, caracterizado por um processo de causalidade, implicaria que não estaríamos diante de duas entidades distintas, mas diante de um único sistema caracterizado por vários níveis. Isso daria margem ao entendimento que estabelece uma identidade entre os *qualia* e os processos cerebrais. No naturalismo biológico searleano, há a admissão de uma redução causal, mas, estranhamente, há também uma redução ontológica das propriedades subjetivas às propriedades objetivas – embora seja veementemente negado por Searle. Portanto, Searle defende uma não-identidade de tipo na relação entre cérebro e consciência.

Segundo Searle (1998), a consciência é um fenômeno essencialmente biológico, ou seja, assume um aspecto biológico ordinário do mundo. Nessa perspectiva, a consciência não se confunde com a cognição, autoconsciência ou conscienciosidade, mas uma espécie de concentração que varia em grau, a depender do nível de sonolência e desatenção. De acordo com o filósofo, os estados mentais apresentam três aspectos essenciais: 1) são fenômenos internos – acontece dentro do cérebro; 2) são qualitativos – cada estado mental requer um modo de senti-lo; 3) são subjetivos – são experienciados por um indivíduo (humano ou animal), ou seja, apresentam-se em uma ontologia de primeira pessoa.

Consoante Dennett (1993), há uma falta de clareza metafísica na ontologia subjetiva searleana, haja vista que ele nos apresenta a estranha categoria da “ontologia epistemológica”, a qual é submetida à consciência. Nessa perspectiva, não seria possível uma subjetividade

epistêmica – seria inadmissível supor que um robô, por exemplo, poderia realizar estados conscientes, já que a subjetividade seria irreduzível a processos físicos. Neste interim, de um lado temos o pensamento searleano cujo caráter é estabelecido por defender a noção de que a consciência é essencialmente subjetiva e irreduzível ao cérebro, do outro, o funcionalismo dennettiano que defende que podemos conceber a relação entre a nossa consciência e a ciência objetiva. Mas a consciência sem o aspecto subjetivo, dirá Searle, se tornaria algo confuso e de caráter contra-intuitivo. Dennett contra-argumenta dizendo que devemos primar por uma explicação transparente do fenômeno, a fim de que não sejamos conduzidos a teorias confusas, permeadas por fantasias calcadas na intuição.

Dennett (1991a;1999), ao tecer sua crítica contra o pensamento searleano, explora a dicotomia ciência e não-ciência – ou pseudociência. Segundo o filósofo, o pensamento searleano é debilitado por enaltecer as noções de consciência da psicologia popular. Dennett considera falacioso o argumento que admite uma consciência subjetiva, ontologicamente irreduzível ao cérebro, como dado epistemológico, ou seja, extraído da ciência objetiva. De acordo com o filósofo, a postura searleana é insustentável, na medida em que apresenta sérias contradições e paradoxos.

Por outro lado, Searle (1998) acusa Dennett de pautar sua crítica em uma teoria exclusivamente objetiva de cunho verificacionista. Isso se reflete na tendência de utilização de métodos objetivos (heterofenomenológicos) em sua investigação – evitando tudo que não possa ser submetido ao rigor da ciência padrão – implicando na postura negacionista a qualquer objeto que fuja do método científico. Dado que a subjetividade é imune a essa metodologia, não é tomada como real (pelo menos até que seja passível da aplicação do método verificacionista). Nesse sentido, o caráter subjetivo da consciência não pode ser examinado, se configurando como irrelevante para as pesquisas sobre a mente. De sorte que Dennett afirma que devemos adotar uma das posições: a intuitiva – a de Searle, que em certo grau é travestida de psicologia popular ou senso comum; a contra-intuitiva – a dennettiana, uma teoria que está em consonância com a ciência objetiva. Sarcasticamente, Dennett conclui afirmando que o defensor do naturalismo biológico deve evitar esse “conjunto de verdades domésticas” e engendrar uma teoria que considere a metodologia da ciência padrão, se quiser que seu projeto de pesquisa seja levado a sério (cf. Searle, 1998, p. 137).

Segundo Searle (2004), apesar da consciência não ser redutível ontologicamente a processos cerebrais, é perfeitamente possível que o mental (a subjetividade), com seus aspectos intrínsecos, seja físico – numa perspectiva mais ampla do termo –, porquanto “[...] a consciência é uma característica do cérebro e, portanto, uma parte do mundo físico” (Searle, 2004, p. 115 –

nossa tradução). Sua proposta é ampliar o conceito de físico, engendrando um que abarque a subjetividade (ontológica) e que evite a alcunha de dualistas. Nesse sentido, a característica essencial da mente (a subjetividade) participaria da realidade física tanto quanto os entes objetivos. Há, portanto, um abandono da noção de que separar o mundo físico e o mental implica em uma irreducibilidade. Essa é a configuração de sua proposta para a solução do problema da causação mental (cf. Prata; Lima Filho, 2013, p. 199). Searle (2004, p. 210 – nossa tradução) nos traz um panorama que facilita mais o entendimento de seu pensamento:

Uma vez que nós abandonamos esse pressuposto, a resposta para os dois enigmas é, primeiramente, que o mental é simplesmente uma característica (no nível do sistema inteiro) da estrutura física do cérebro e, em segundo lugar, causalmente falando não existem dois fenômenos independentes, o esforço consciente e as descargas neuronais inconscientes. Existe apenas o sistema cerebral, que tem um nível de descrição onde descargas neuronais estão ocorrendo e outro nível de descrição, o nível do sistema, onde o sistema é consciente e de fato tenta conscientemente levantar o braço.

Nessa passagem podemos perceber que a noção de irreducibilidade, apesar de defender uma perspectiva de não-identidade entre as propriedades subjetivas e objetivas, não nega que pode haver uma conexão, uma relação de dependência mútua, dentro do domínio físico (cf. Searle, 1992, p. 117). No entanto, no pensamento searleano, temos uma noção muito rasa de conexão – em um nível que se aproxima de um dualismo – o que não faz com que Searle abandone a convicção de que “tudo o que existe são partículas em campos de força” (Searle, 1992, p. 28), ou seja, há uma generalidade do mundo físico. No que tange a natureza biológica da mente, defende que se trata de um tipo de conexão tonificada, que possibilita evitar o rótulo de dualismo e, ao mesmo tempo, delinear um fisicalismo aceitável. Eis a réplica de Dennett (1991a, p. 279 – nossa tradução):

Muitos filósofos, particularmente aqueles influenciados pela escola Husserliana de fenomenologia (Dreyfus, 1979; Searle, 1983), enfatizaram a importância desse “pano de fundo” da experiência consciente, mas eles normalmente o descreveram como uma característica misteriosa ou recalcitrante, desafiando a explicação mecânica, ao invés da chave [...] para fornecer uma teoria computacional do que acontece. Esses filósofos supuseram que a consciência é a fonte de algum tipo especial de “intencionalidade intrínseca”, mas, como observou o filósofo Robert van Gulick, isso leva a contradição.

De acordo com Dennett (1991a), os estados mentais, a experiência consciente não deve ser concebida nos parâmetros do dualismo, ou seja, a teoria que fundamenta o “teatro cartesiano” se mostra inviável justamente por considerar a mente de modo ontológico. Não há no cérebro uma localização específica segundo a qual a consciência desenvolve seus processos. A simples admissão dessa possibilidade nos leva à falaciosa noção de que a intencionalidade é intrínseca à consciência.

Sem embargos, Searle (1998) está convicto de que nossos estados mentais são causados por processos cerebrais de nível inferior. Esses processos ocorreriam nos neurônios, em níveis

micro das sinapses, grupos de neurônios e em complexos celulares. Considera, portanto, os trabalhos das neurociências para tomar a consciência como um fenômeno essencialmente biológico. No entanto, o filósofo entende que isso não nos dá o direito de conceber o cérebro como um computador, ou seja, há pouca fidelidade na analogia que envolve cérebros e computadores. Assim, com o intento de deflagrar uma crítica contundente contra a *inteligência artificial forte*<sup>96</sup>, o filósofo vai elaborar um dos mais famosos experimentos de pensamento filosófico, a saber: o argumento do quarto chinês.

Em consonância com Cooper (2005), a elaboração de um experimento de pensamento é bem-sucedido quando se tem um domínio dos conceitos subjacentes. Nesse sentido, antes de examinarmos o EP searleano é preciso um maior entendimento dos três níveis básicos da linguagem – já que esses serão fundamentais para uma boa compreensão do EP em questão: a primeira seria a sintaxe, cujo papel na estrutura de uma língua se resume em uma relação formal entre os signos, ou seja, considera as regras que constituem a linguagem – quando se utiliza o sinal indicativo de crase, por exemplo, para indicar a junção da preposição “a” com o artigo definido “a” se está estabelecendo o uso da norma culta da língua portuguesa; a segunda seria a semântica que tem por finalidade estabelecer a relação do termo com o mundo de forma literal – uma tentativa de fazer com que a palavra aponte para o objeto ou fato que ela quer significar; e por fim, a pragmática, que examina o uso da linguagem entre indivíduos, ou seja, a partir de uma interlocução se espera determinado comportamento. Agora estamos aptos a compreender melhor o experimento mental de Searle.

Considere um sujeito que domina exclusivamente a língua inglesa e que se encontra enclausurado em uma sala. Nesse confinamento, ele tem acesso a três cartapácios específicos: um com textos em chinês (o roteiro ou programa); outro com textos em sua língua nativa, o inglês, que lhe permite realizar correlações com os dois textos (histórias ou perguntas); e o terceiro possibilita uma articulação com os dois primeiros – os nativos de língua chinesa são os que têm acesso a esse cartapácio. Ao ser introduzida na sala uma pergunta, por exemplo, o confinado irá consultar o roteiro (o programa) e externalizar a resposta mais adequada. Nesse sentido, os falantes nativos da língua chinesa irão conceber que estão diante de um genuíno falante da língua chinesa. Todavia, não houve nesse processo a articulação dos três níveis básicos da linguagem, mas tão-somente uma aplicação das normas sintáticas, ou seja, uma

---

<sup>96</sup> Também denominada de *conexionista* – uma simulação do cérebro, baseada em redes neurais artificiais. (cf. Tripicchio, 2004, p. 86).

manipulação de signos. O confinado não sabe chinês. Ele apenas se comporta como soubesse (cf. Searle, 1998, p. 37).

Com efeito, por definição, um computador é um conjunto de símbolos que pode ser organizado de modo a ser interpretável, ou seja, um computador exerce uma função sintática. Esses símbolos têm seus significados atribuídos por um observador (uma mente), de tal forma que esse *status* que configura o sistema não é intrínseco ao próprio computador – mundo físico. O computacionalista tenta explicar a existência de um observador pela computação. Admite que o cérebro é um computador para fundamentar a ontologia da mente, mas se a computação – sistema simbólico – é atribuída por um observador, ela não pode ser a explicação da existência do observado. Dito de outro modo, a explicação do fato dos seres humanos serem conscientes não pode vir da hipótese que afirma que o cérebro é um computador. Esse argumento é uma tentativa de estabelecer uma falsa analogia.

A crítica searleana à IA forte está clara. Uma mente humana utiliza os três níveis básicos da linguagem na sua atividade. Um computador é caracterizado exclusivamente pelo uso de um programa que processa seus dados considerando tão-somente o aspecto sintático da linguagem. Seu comportamento, nesse sentido, seria um simulacro, pois não há um processo cognitivo. O argumento de Searle (1998, p. 38) para refutar a IA forte se dá da seguinte maneira: 1) programas são totalmente sintáticos; 2) as mentes têm uma capacidade semântica; 3) a sintaxe não é a mesma coisa que a semântica, nem é por si só suficiente para garantir um conteúdo semântico. Consequentemente, programas não são mentes.

Havendo a possibilidade de refutação de uma dessas premissas, a tese da identidade entre mentes humanas e computadores pode ser revista. Entrementes, examinemos a estrutura do argumento com mais atenção. A primeira premissa é corroborada pela definição de Alan Turing. Nossas mentes têm a capacidade de alcançar os significados dos termos de uma língua, é o que diz a segunda premissa. A terceira, por fim, defende que o domínio sintático é insuficiente para obter a compreensão dos significados e referência dos termos de uma língua. Os computadores não executam processos semânticos, a consciência sim, portanto, a mente não é análoga a um programa de computador, a um *software*, ou seja, “a mente não poderia ser apenas um programa de computador, já que os símbolos formais dos programas de computador, tomados isoladamente, não são suficientes para garantir a presença do conteúdo semântico que ocorre na mente” (Searle, 1998, p. 38).

O experimento de pensamento delineado por Searle funciona bem no contexto da sua argumentação contra a IA forte. Seu êxito está no uso de conceitos bem definidos, o que não deixa margem para ambiguidade – estaria em consonância com a tese browniana dos

experimentos de pensamento por não violar nenhuma lei da natureza. No entanto, o filósofo parece ecoar Descartes quando nega a possibilidade de um *software* tornar um robô consciente. Seu experimento de pensamento procura dissuadir o leitor a considerar essa possibilidade (cf. Dennett, 1991a, p. 440). Eis a réplica de Dennett (1991a, p. 435 – nossa tradução):

Como o caso de Mary de Frank Jackson, a cientista das cores, o experimento mental de Searle produz uma convicção forte e clara apenas quando deixamos de seguir as instruções. Essas bombas de intuição estão com defeito; elas não aumentam, mas enganam nossa imaginação.

De acordo com Teixeira (2008), grande parte dos antifuncionalistas concentram a sua rejeição na acusação de que Dennett edificou uma teoria da mente que não alcança a ontologia da consciência. Searle é seu representante maior. Segundo eles, a tese dennettiana se resume a uma simplificação tosca da consciência, ao ignorá-la intencionalmente. Ela não descreve os processos físicos que dão origem a experiência subjetiva, tampouco explica sua natureza. Nessa perspectiva, Dennett se resume a estabelecer uma correlação entre os processos físicos e a emanção da consciência. Essa crítica poderia ser rebatida em uma acusação que revela a exploração searleana do aspecto polissêmico do termo consciência – uma postura que abre precedente para várias versões do que seria a consciência. Nesse sentido, Searle parece oscilar entre um fisicalismo e um dualismo, porquanto, como Chalmers, parte de uma noção de consciência e, a partir daí, procura construir uma crítica (cf. Teixeira, 2008, p. 30).

Uma das principais características da teoria da consciência dennettiana é a tentativa de identificar e demonstrar a possibilidade de superar essa oscilação, esse movimento pendular entre fisicalismo e dualismo que tomou conta das discussões sobre a ontologia da inteligência consciente, oferecendo uma resolução inovadora para o problema – já que dualistas e materialistas não logram êxito na procura dos termos teóricos fundamentais que podem lhes proporcionar a descrição fiel da coisa-mente.

Outro ponto da teoria dennettiana, duramente criticada por Searle (1998), é o conceito ou noção de meme. Como vimos no capítulo II, o meme é uma versão do gene (conceito que ajuda a explicar a evolução biológica darwiniana) para descrever a evolução cultural, ou seja, todos os aspectos que envolvem nossa cultura. Nesse sentido, a teoria memética vislumbra a explicação de todos os aspectos da cultura humana – que por vezes relacionamos com o espírito criativo do homem – de forma naturalista. Podemos, por exemplo, desenvolver explicações sobre as religiões, a arte e a ética a partir da noção de um conceito que assume as características que explicam a evolução biológica, o desenvolvimento das espécies. Não obstante, o filósofo discorda dessa analogia, dada a qualidade diferente de suas manifestações. A função dos genes

na evolução biológica, segundo o filósofo, é fruto de resultados de forças naturais brutas e obscuras. Para o autor, essa característica não é observada na noção de meme.

Na cultura humana, de acordo com Searle (1998), o compartilhamento de ideias e teorias por meio da “imitação” é realizada em uma economia que envolve processos de seleção e direcionamento consciente, ou seja, exige escolhas conscientes de itens para uma finalidade específica. Nesse sentido, enquanto o gene proporciona a evolução por processo estritamente não consciente, o meme (entendido como um termo que remete à imitação) não tem essa característica, ou seja, requer uma escolha consciente daquilo que deve ser perpetuado. Portanto, há a necessidade de compreensão e interpretação das ideias, bem como a implicação de uma certa noção de intencionalidade. Por conseguinte, fica evidente que nem todas as ideias que emergem em nossa cultura serão necessariamente repassadas para a geração futura, pois são submetidas a um processo de julgamento e de teste. Antes disso, não podem se tornar “candidatas” à rejeição ou imitação. Searle (1998, p. 125) continua em sua análise acerca do meme argumentando sobre este que:

a difusão de ideias através da imitação requer todo o aparato da consciência humana e intencionalidade. As ideias devem ser compreendidas e interpretadas. Só depois de entendidas e julgadas desejáveis ou indesejáveis elas podem ser tratadas como candidatas à imitação ou rejeição. A imitação requer tipicamente um esforço consciente da parte do imitador. O processo todo normalmente envolve a linguagem com a sua variabilidade e sutileza. [...] a transmissão de ideias através da imitação é totalmente diferente da transmissão de genes através da reprodução. Logo, a analogia entre gene e memes é enganosa desde o início.

Searle (1998) nega que o meme possua as características análogas a dos genes e está convicto que essa teoria é falaciosa em sua essência, haja vista que a transmissão cultural das ideias é realizada por meio de seleção consciente – enquanto que os genes realizam sua transmissão por meio da reprodução.

O processo de evolução cultural, dirá Dennett (1991a; 1995a; 1998), não se dá exclusivamente a partir da atividade de seleção consciente, haja vista que não se observa tal procedimento na construção de narrativas. Além disso, a neurociência parece corroborar a tese memética, ao apresentar os resultados das pesquisas sobre neurônios-espelho. A rigor, não faz sentido afirmar que esses grupos de neurônios apresentam consciência, pois estaríamos voltando ao problema dos homúnculos.

De acordo com Dennett (1991a), as críticas à noção de memes estão apoiadas em dois pontos principais: 1) a evolução cultural não é análoga à evolução darwiniana, mas sim, à lamarckiana; 2) a evolução cultural apresenta como característica a transmissão horizontal, ou seja, alguns elementos não são limitados pelo aspecto hereditário – diferentemente da evolução darwiniana, que são livres para saltar de linhagem para linhagem. Assim, dado que duas

espécies animais – peixes e anfíbios, por exemplo – apresentam linhas evolutivas separadas, as inovações de um não podem ser adquiridas pelo outro. Esse privilégio será de seus descendentes diretos, mas isso não é observado na evolução cultural. No entanto, subjacente a essas acusações está a ideia de que a reprodução sexual é a responsável direta pelo aprimoramento das espécies na evolução darwiniana.

Com efeito, a evolução darwiniana não apresenta a reprodução sexual como um elemento estritamente obrigatório. Um número considerável – 99% de toda evolução darwiniana em nosso planeta – de aprimoramentos de espécies vivas se deu entre replicadores que se reproduzem de modo assexuado. Além disso, a própria reprodução sexual é uma característica evoluída e não uma pré-condição para a evolução darwiniana. Por conseguinte, a ausência do aspecto reprodutivo de cunho sexual na tese memética não é um desafio que deva ser considerado como limitador. Esse argumento pode apresentar maior eficácia se for direcionado às “supostas *desanalogias* do lamarckianismo e transmissão horizontal ou anastomose – união de linhagem” (Dennett, 1998, p. 08 – grifo nosso).

A validade científica da consciência, aparentemente, está no epicentro da discussão entre Dennett e Searle. Há uma exploração searleana dos conceitos da psicologia popular, dirá Dennett, e isso enfraquece o naturalismo biológico. Todavia, Searle admite essa exploração e entende que isso não é motivo para rejeitar sua tese – na verdade essa crítica, segundo o filósofo, se configura como um preconceito e não como um argumento –, porquanto, uma abordagem calcada na psicologia popular não exclui as descobertas da ciência padrão. Dennett acusa Searle e os demais críticos de anacronismo em seus argumentos, pois, à medida que nutrem uma esperança de solução do problema da consciência, especulam que ela virá da análise do mundo físico, mas não da biologia ou da ciência cognitiva.

### **3.3 O Self: o centro de gravidade narrativa**

Temos a inclinação natural de identificar, rotular e conceituar as entidades com as quais temos contato. Essa postura nos conduz a elaborar explicações que nos permitem realizar previsões e controlar alguns fatos acerca do mundo e também de nossas mentes. Mas o que aconteceria se essa mesma inclinação nos conduzisse a imaginar coisas e causas que não fazem parte da realidade física? Cada acontecimento no mundo poderia ser explicado a partir de formulações de hipóteses com um alto teor de misticismo. Segundo o pensamento dennettiano, a maioria das teorias sobre o *self* (o eu) assume esse caráter. O que fazer se somos coagidos a procurar a “causa” que causa tudo? Como nomear esse algo? Enquanto o chamo de “você”, você chama de “eu”. Nesta seção iremos examinar a concepção narrativista do *self* como

solução para o problema da consciência de identidade. Antes de aprofundarmos nossa análise, vejamos o que dizem as pesquisas da neurociência sobre o nosso sentimento de identidade.

Sentimo-nos confortavelmente aprisionados em nosso corpo. Todavia, a segurança que temos, o sentimento de posse que estabelecemos na relação com nosso corpo, são suscetíveis a falibilidade e a flexibilidade. Essa possibilidade nos é esclarecida por Ramachandran (2014, p. 316) ao nos afirmar que:

acredite ou não, mediante um truque óptico podemos ser induzidos a deixar temporariamente nosso corpo e nos experimentar em outro lugar [...] usando uma maquiagem pesada para nos disfarçar e olhando para nossa própria imagem num vídeo (que não precisa fazer uma inversão esquerda-direita como no espelho), podemos ter uma ligeira ideia de uma experiência de saída do próprio corpo.

Nesse sentido, a neurociência consegue reunir evidências que nos ajudam a compreender o funcionamento do cérebro e nos permite realizar uma abordagem mais científica sobre o *self*. A investigação de alguns distúrbios ocupa papel fundamental nessa empreitada, pois cada um deles “torna-se uma janela para um aspecto do *self*” (Ramachandran, 2014, p. 315).

De acordo com Ramachandran (2014), podemos elencar alguns aspectos que envolvem nosso sentimento de identidade (aspectos que são, num primeiro momento, incontestáveis, haja vista que os acessamos intuitivamente): *unidade* – as experiências sensoriais e emocionais são reunidas em um único bojo a formar um sujeito; *continuidade* – embora uma diversidade de eventos seja uma constante em nossa vida, há um sentimento de continuidade e de identidade; *incorporação*<sup>97</sup> – nos sentimos “presos” e ao mesmo tempo a vontade em nosso corpo; *privacidade* – temos acesso exclusivo a nossos *qualia* e a nossa vida mental; *inserção social* – nossos sentimentos (ciúme, raiva, orgulho, humildade, etc.) não teria significado fora de uma interação social, ou seja, nutrimos um soberbo senso de privacidade e autonomia, em um contraponto com outros cérebros; *livre-arbítrio* – construímos uma representação de nós mesmos como responsáveis e senhores de nossas escolhas; por fim, temos a intuição da *autoconsciência* o que, de acordo com o cientista, possivelmente “dependa em parte de nosso uso de neurônios-espelho de uma maneira recursiva, permitindo-nos ver a nós mesmos a partir do ponto de vista de outra pessoa” (Ramachandran, 2014, p. 318).

---

<sup>97</sup> Há pelo menos três distúrbios que caracterizam a relevância da incorporação para a constituição de nosso *self*, a saber: *apontemnofilia* – um estranho desejo de amputar um dos membros do corpo – um sentimento inefável que leva o sujeito a rejeitar uma parte de seu corpo como se não o constituísse – que, segundo Ramachandran, está relacionado a uma anomalia no hemisfério direito do cérebro; a *transexualidade* – uma sensação de incompatibilidade psicossomática em que o *self* (que aparentemente possui um gênero) não se reconhece no gênero do corpo que habita; e a *somatoparafrenia* – em uma pequena parte dos pacientes de acidentes vasculares cerebrais apresentam uma rejeição a um dos membros paralisados, o atribuindo a uma pessoa do círculo familiar, um dos pais, irmãos ou cônjuge (cf. Ramachandran, 2014, p. 320-26).

Consoante Dennett (1991a), há duas possíveis explicações sobre o porquê de nossa fácil sedução em relação a essas ideias: 1) o modo pelo qual mantemos a ideia de necessidade de processos apartados de observações internas que intervêm em momentos que devemos relatar, bem como do relatório que emitimos – de tal modo que acabamos por negligenciar “o fato de que em algum ponto este retrocesso de observadores interiores deve ser interrompido por um processo que une os conteúdos a sua expressão verbal, sem nenhum apreciador de conteúdo intermediário” (Dennett, 1991a, p. 319); 2) as conexões de nossa rede de comunicação internas, esculpida desse modo pela evolução natural, possuem de fato a prerrogativa de “organizar nossas mentes em sistemas reflexivos ou de automonitoramento indefinidamente poderosos” (*Ibidem*). O poder reflexivo da consciência, justificadamente segundo o filósofo, é identificado como o seu âmago. Há a possibilidade de recorrermos à psicologia popular para uma descrição dos sistemas de automonitoramento, não obstante, esse recurso não estava isento de recair em um “materialismo” cartesiano. O autor tenta obliterar essa possibilidade a todo custo (cf. Dennett, 1991a, p. 319-20).

Malgrado, tenhamos, até agora, abordado os mecanismos subjacentes ao funcionamento da consciência em um nível subpessoal da teoria da mente dennettiana – nível que pode ser revisado à medida que surgirem novas descobertas nas ciências empíricas – ainda nos parece crucial examinar o problema do nível do agente pessoal. Esse problema é considerado por muitos teoricamente estável e filosoficamente orientado – a despeito da periódica aproximação com as ciências cognitivas. De modo que define um método de abordagem do *self* calcado em um argumento que é similar ao exame que identifica a causa de nossas dores. De forma bastante eloquente, Dennett, (1969, p. 93 – nossa tradução) contribui conosco:

Não localizamos nossas dores com o auxílio de quaisquer qualidades ou "sinais locais" descritivos de forma independente, fornecidos pelas sensações; [...] O que quer que o cérebro esteja 'fazendo' quando localiza uma dor, a pessoa não faz nada no processo de localização de suas dores, pois não existe esse processo em que uma pessoa possa se envolver. Alguém poderia se envolver no processo localizando as dores de outra pessoa, fazendo perguntas, cutucando até que ele grite e assim por diante, mas não no processo de localizar as próprias dores.

Sem embargo, na esteira de Dennett (1969), nosso exame perpassa a questão dos níveis subpessoais<sup>98</sup> e das narrativas conscientes, alcançando a problemática que envolve o modo pelo qual adquirimos autoconsciência de nós mesmos enquanto sujeitos no mundo, ou seja, como estabelecemos uma distinção entre o “eu” e o ambiente. Assim, segundo o filósofo, do desvelamento das capacidades do nosso cérebro e da concepção da consciência como

---

<sup>98</sup> A discussão sobre a distinção entre os níveis subpessoal e pessoal é melhor abordada em *Content and Consciousness*, se comparada com *Consciousness Explained*.

propriedade fenomênica, das edições de narrativas realizadas por essa máquina, emerge um centro gravitacional<sup>99</sup>. Nesse sentido, delineamos nossa consciência de identidade a partir de uma unificação, um sujeito que age por conta própria e que possui um corpo ao passo que realiza um comando virtual acima dos níveis dos rascunhos múltiplos, obedecendo os processos pandemônicos. Esse comando virtual envolve um *self*, um “eu” que é extraído de modo abstrato das narrativas autobiográficas engendradas por arquivos de memória<sup>100</sup> em nossa máquina joyceana – e em nosso cérebro –, algo similar ao que ocorre na física quando se estabelece um “ponto abstrato a partir das forças de atração exercidas por cada partícula de um corpo, podendo produzir assim posteriormente uma representação de si com base linguística” (De Paulo, 2018, p. 51). Uma questão que emerge desse argumento é se não seria possível que essas características que envolvem a unidade e o controle não passassem de ilusões elaboradas pelo *self*, ou, pior ainda, o *self* é que poderia ser uma ilusão se tomado como um ente que pode ser reduzido a outros tipos de entidades.

No modelo apresentado por Dennett (1991a), o *self* é assimilado como um centro virtual que envolve o processamento de informações subpessoal e o apercebimento. O surgimento da razão subvenciou o limiar entre o “eu”, o ambiente e os outros seres. O processo de auto-identidade, segundo o filósofo, é observado nos mais simples seres vivos. Essa atitude – expressa em comportamentos como se proteger, por exemplo – percebida em várias espécies são os “eus” biológicos e esses “eus” simples não seriam algo concreto, mas um princípio abstrato de organização. São de caráter permeável e indefinido, um “exemplo da mãe natureza tolerando o erro se o custo for justo” (Dennett, 1991a, p. 414 – grifo nosso).

Com efeito, nosso cérebro recorre a uma estratégia – que garante autocontrole, autodefinição e autoproteção – calcada nas elaborações de narrativas. Em um nível particular, criamos e manipulamos estórias sobre nós mesmos com o intuito de compartilharmos com outros indivíduos nossa identidade (sem abrir mão de realizar uma representação para nós mesmos. As aranhas e os castores não precisam refletir como executar suas tarefas peculiares (tecer teias e edificar represas), não há deliberação em seus trabalhos individuais. Por outro lado, as tecelãs e os engenheiros necessitam de uma narrativa deliberada para descrever suas

---

<sup>99</sup> A concepção dennettiana de identidade narrativa difere daquela delineada por Paul Ricoeur (1913-2005), segundo a qual as ações praticadas ao longo do tempo, uma narratividade incoativa que permite ao indivíduo uma descrição sobre si mesmo (cf. Reichert, 2019, p. 51), mas que não descreve a origem da composição dessas narrativas ou os processos subjacentes à atividade cerebral que lhes constituem.

<sup>100</sup> A memória autobiográfica pode ser definida como a capacidade que o cérebro tem de repressar as experiências e o conhecimento referente ao *self*. São caracterizadas por obedecer a uma transitoriedade por um dinamismo em sua edificação calcada em uma série de acontecimentos vivenciados pelo indivíduo (cf. Baddeley, 2011, p. 153).

especializações e para se autodescrever. Nossas histórias tendem a cristalizar nossa identidade, pois é mais comum que sejamos influenciados por elas do que o contrário. A consciência humana – nossa identidade narrativa – não é sua fonte, mas seu produto. De forma bastante eloquente, Dennett (1991a, p. 418 – nossa tradução) contribui conosco:

Essas sequências ou fluxos de narrativa surgem como se fossem de uma única fonte – não apenas no sentido físico óbvio de fluir de apenas uma boca, ou um lápis ou caneta, mas em um sentido mais sutil: seu efeito em qualquer público é encorajar para (tentar) postular um agente unificado cujas palavras são, sobre quem são: em suma, para postular um centro de gravidade narrativa.

De acordo com Dennett (1991a), apenas o gênero humano tem a capacidade de construir “eus” (*selves*) narrativos – calcados em teias da linguagem –, porém, há criaturas biológicas portadoras desse tipo de categoria pré-linguística. Esse tipo de categoria lhes permite delinear limites-de-si cujo papel é o de estabelecer sua distinção com o ambiente. Nesse sentido, infere-se que, sendo essa categoria anterior ao eu, o *self* se resume a uma narrativa, uma espécie de extensão linguística do *self* biológico segundo o qual possibilita unidade<sup>101</sup> e autopreservação por intermédio de narrativas autobiográficas cujo sistema realiza não só uma autorrepresentação, mas também uma representação para outras pessoas. Dennett (1991a, p. 417 – nossa tradução) assevera – e nos explica melhor:

Nosso ambiente humano contém não apenas comida e abrigo, inimigos para lutar ou fugir e específicos com quem acasalar, mas também palavras. Essas palavras são elementos poderosos de nosso ambiente que prontamente incorporamos, ingerindo e extrusando-as, tecendo-as como teias de aranha em fios narrativos autoprotetores. Na verdade [...] quando permitimos a entrada dessas palavras, esses veículos-meme, elas tendem a assumir o controle, criando-nos a partir da matéria-prima que encontram em nosso cérebro.

Com efeito, podemos concordar com Dennett (1991a) que o *self* não é a origem de narrativas, mas seu produto. Isso nos permite lhe atribuir o rótulo de “centro de gravidade” abstrato consolidado a partir de um conjunto de narrativas, dada a urgência de autopreservação e de autocontrole. O filósofo descarta a possibilidade de uma autorrepresentação primitiva, ou seja, uma representação essencial que permita delinear a consciência de si de modo apriorístico – de tal forma que o *self* é considerado sem os processos de informação que são característicos ao modelo pandemônico acompanhado por narrativas conscientes do cérebro. Segundo o autor, o *self* seria esse centro de gravidade abstrato que simplesmente não possui uma precisão espaço-temporal, sendo concebido como uma personagem construída de acordo com as narrativas elaboradas por nosso cérebro. A abstração aqui descrita se aproxima a um roteiro em que uma das personagens é inserida no decorrer da história e em seguida desenvolvida. Por conseguinte,

---

<sup>101</sup> Em Dennett (1991a;1969) essa unidade não seria necessariamente ontológica e temporalmente real.

o *self* não seria nada mais do que uma ficção útil produzida sem processamentos inteligentes e conscientes de informação, calcada em rascunhos múltiplos e vitórias de conteúdos que seguem padrões algorítmicos – isenta da autonomia do sujeito. Portanto, descarta-se a possibilidade, no pensamento dennettiano, do *self* emergir de uma ação voluntária (cf. Dennett, 1991a, p. 418).

Não há contradição, portanto, em afirmar que um ente absolutamente abstrato como o *self* tem sua origem em uma entidade concreta e consciente. Decerto, é perfeitamente possível esse ente abstrato ser o resultado de atividades de um ente concreto e ignorante, envolvendo processos neurofisiológicos automáticos de informação, como os que ocorrem no cérebro humano. Nesse sentido, o *self* pode ser entendido como um centro de controle operacional indeterminado que auxilia na manutenção de um comportamento racional e expressão de unidade cuja identidade pessoal é emanada. Dennett (1991a, p. 418 – grifo nosso) assevera – e nos explica melhor:

Os físicos apreciam a enorme simplificação que você obtém quando postula um centro de gravidade para um objeto, um único ponto em relação ao qual todas as forças gravitacionais podem ser calculadas. Nós, heterofenomenólogos, apreciamos a enorme simplificação que você obtém quando postula um centro de gravidade narrativa para um corpo humano girando em narrativa. Como o *self* biológico, este *self* psicológico ou narrativo é mais uma abstração, não uma coisa no cérebro, mas ainda um atrator de propriedades notavelmente robusto e quase tangível, o "proprietário do registro" de quaisquer itens e características que estejam mentindo sobre não reclamados. Quem é o dono do seu carro? Você faz. Quem é o dono de suas roupas? Você faz. Então quem é o dono do seu corpo? Você faz!

Nesse sentido, o “eu” é construído a partir de uma representação de nós mesmos, calcada em uma seleção de semi-narrativas que, por sua vez, é filtrada em outras mini-narrativas, e outras ainda menores – todas elas formadas a partir de um banco de dados angariados da nossa cultura (o memoplex). Essa triagem de versões editadas, que culmina na representação do “eu”, é o centro de gravidade narrativa – um afunilamento de versões de nós mesmos nos parâmetros de uma seleção natural. A figura 3.1 ilustra as fases do processamento da construção do self.

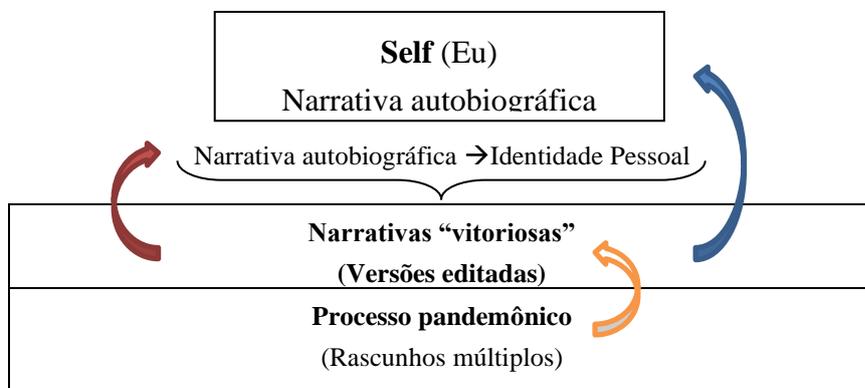


Figura 3.1

### Fases<sup>102</sup> do processo de construção do *self*

Por conseguinte, aquilo que chamamos de “eu” não é uma criação miraculosa, nem supõe a existência independente de uma alma, mas um produto de processos sociais que nos caracterizam, passíveis de mudanças abruptas. A rigor, o *self* adquire uma estabilidade a partir do acúmulo de crenças que o constitui. Ao prescrever, essas crenças promovem, de modo permanente ou temporariamente, um sujeito caduco levando a necessidade de uma nova construção narrativa sobre si, uma atualização do *self* (cf. Dennett, 1991a, p. 423).

Consoante Dennett (1991a), não há motivos para crises existenciais. Você existe de fato. No entanto, sua existência não é prolongada indefinidamente pelo seu corpo – embora seja uma pré-condição para sua existência. Se considerarmos que somos uma alma, um milagre de caráter imaterial, seria possível descrever nossa imortalidade, mas seria uma entidade de caráter inescrutável – uma *virtus dormitiva*. E se fossemos realmente uma composição que engendra uma substância material, um arranjo sublime de átomos em nosso cérebro, nossa imortalidade se estabeleceria em função das leis da física que fundamentam essa união – o que seria uma “meia-vida” de um eu? Dennett pergunta a um físico. Se considerarmos a tese do centro de gravidade narrativa, por outro lado, veremos que nossa existência está atrelada a perseverança dessa narrativa<sup>103</sup>, que teria o poder, em tese, de permanecer indefinidamente perante as várias mudanças do ambiente e armazenada indefinidamente como informação pura. Dennett (1991a, p. 430 – grifo nosso) afirma de forma clara e vigorosa:

Se você é aquela organização de informações que estruturou o sistema de controle do seu corpo (ou, para colocá-lo em sua forma provocativa mais usual, se você é o programa que roda no computador do seu cérebro), então você poderia, em princípio, sobreviver a morte do seu corpo intacto como um programa pode sobreviver à destruição do computador no qual foi criado e executado pela primeira vez. Alguns pensadores (por exemplo, Penrose, 1989) consideram isso uma implicação terrível e profundamente contra-intuitiva da visão que defendi aqui. Mas se é a imortalidade potencial que você deseja, as alternativas são simplesmente indefensáveis.

Uma possível ofensiva à teoria do *self* dennettiano seria a concretude da patologia do transtorno dissociativo de identidade. O filme *Fragmentado*, de 2016, do diretor Manoj N. Shyamalan, nos apresenta esse distúrbio que chega ao nível do absurdo e da descrença. No longa-metragem, o ator James McAvoy interpreta um homem que apresenta pelo menos vinte e quatro personalidades diferentes, das quais merecem destaques: Patrícia – uma mulher inteligente, meticulosa e autoritária; Hedwig – uma criança que nunca deixa de ter nove anos; Jade – uma possível adolescente (não fica claro no filme) que acredita ter diabetes; Kevin

---

<sup>102</sup> Esquema do autor.

<sup>103</sup> Algo similar aos contos das “Mil e uma noites”, dirá Dennett (1991a), mas todos contados em uma única vez.

Wendell Crumb – a personalidade “originária”. Segundo o roteiro do filme, Kevin criou as outras personalidades, a fim de que pudesse fugir das agressões físicas de sua mãe, mas perdeu o controle sobre as outras personalidades. Embora essa patologia<sup>104</sup> seja um fato, Dennett não acredita que poderia existir – como no filme em questão – vinte e quatro *selves* confabulando. O autor realiza um comentário que aborda essa questão:

Essas crianças muitas vezes foram mantidas em circunstâncias tão extraordinariamente aterrorizantes e confusas que fico mais surpreso por elas sobreviverem psicologicamente do que conseguirem se preservar por meio de um redesenho desesperado de seus limites. O que elas fazem, quando confrontadas com conflito e dor avassaladoras, é o seguinte: elas "vão embora". Elas criam um limite para que o horror não aconteça com elas; isso não acontece a ninguém, ou a alguma outra pessoa, mais capaz de sustentar sua organização sob tal ataque – pelo menos é o que elas dizem que fizeram, da melhor maneira que se lembram (Dennett, 1991a, p. 420 – nossa tradução).

Além do transtorno dissociativo de identidade, segundo Dennett (1991a), existem casos de epilepsia em que é preciso realizar uma intervenção cirúrgica para “separar” os dois hemisférios do cérebro e, como possível sequela, há relatos de casos de dupla personalidade<sup>105</sup>. O filósofo nega que esse fato possa ser uma refutação plausível para sua teoria do *self*, dado que não é “logicamente possível que exista tal *self* do hemisfério direito em um paciente com cérebro dividido, pois minha teoria diz que não existe, e diz o porquê: as condições para acumular o tipo de riqueza narrativa (e independência) que constitui um eu *plenamente desenvolvido* não estão presentes” (Dennett, 1991a, p. 426 – nossa tradução). Um *self* não seria um ponto matemático qualquer. De acordo com o autor, é uma abstração estabelecida pela vasta rede de (auto) atribuições e (auto) interpretações que fizeram e fazem parte de um organismo cujo centro de gravidade narrativa é a sua culminância. Em termos concretos, isso significa que o *self* desenvolve um trabalho de extrema relevância “na economia cognitiva contínua desse corpo vivo, porque, de todas as coisas no ambiente que um corpo ativo deve fazer são os modelos mentais, e nenhuma é mais crucial do que o modelo que o agente tem de si mesmo” (*Ibidem*). Em que lugar nos deparamos com a nossa autorrepresentação? Onde você estiver, dirá Dennett. O que isso significa? Nada além do que seu centro de gravidade narrativa (cf. Dennett, 1991a, p. 429).

A teoria do centro de gravidade narrativa emerge como uma ferramenta importante para os enigmas dos transtornos de personalidade múltipla e transtornos de personalidade fracionada, que apresentam cessações e processamentos heterogêneos dessas narrativas que

---

<sup>104</sup> A existência do transtorno dissociativo de identidade é uma questão controversa.

<sup>105</sup> Alguns casos semelhantes foram analisados por Derek Parfit (1942-2017), em um artigo, de 1983, intitulado *Personal Identity*.

podem reestruturar os *selves* de pacientes acometidos desse fenômeno. Podemos descrever os *selves* como clubes cuja organização se dá por meio de uma legislação – podem ser dissolvidos e reintegrados no processamento contínuo do cérebro. Como o que vai preservar os aspectos originais do clube serão a permanência do comportamento característico de seus membros a partir da observação das suas diretrizes, analogamente, os *selves* observará suas regras gerais de construção de narrativas, refletidas na elaboração e aproveitamento de semi-selves, quase-selves e selves transitivos (cf. De Paulo, 2018, p. 54).

Em Dennett (1991a; 1991b), podemos perceber que apesar do cérebro engendrar uma única versão do “eu”, é perfeitamente plausível que essa versão seja constantemente submetida a reconstruções e alterações. Questionar se o nosso cérebro edifica uma única ou várias versões do “eu” é totalmente pertinente, haja vista que é comum nutrirmos dúvidas concernentes a determinados aspectos que nos constituem (projetos, nossos verdadeiros talentos, etc.)

O *self* de que nos fala Dennett (1991a) é totalmente passível de patologias como a observada no filme de Shyamalan. Sua teoria contribui para as pesquisas que envolvem não só o transtorno dissociativo de identidade<sup>106</sup>, mas também de outros que abarcam a perda da noção da existência de um “eu” ou mesmo de um único “eu”. Ela não postula um *self* primitivo *a priori*. Segundo o filósofo, um “eu” como resultado de narrativas está muito mais vulnerável a acidentes fatídicos e ser maculado por distúrbios (cérebro dividido, esquizofrenia, paranoia, autoengano, etc.).

Destarte, Dennett (1991a) nos apresenta uma teoria que vislumbra um modelo de rascunhos múltiplos, que concebe a mente como um pandemônio e a consciência como uma espécie de máquina Joyceana virtual. Essa abordagem permite esclarecer algumas patologias que assombram a psicologia e intrigam a neurociência em geral. No entanto, ela faz ressurgir um problema antigo da filosofia: até que ponto podemos atribuir imputabilidade moral<sup>107</sup> a um *eu* narrativo<sup>108</sup>? Em termos concretos, há uma colisão entre a natureza ficcional e funcional do *self* com o sujeito tradicional – portador de vontade livre. O filósofo responde a essa questão explicando que o livre arbítrio também foi esculpido pela evolução natural, ou seja, a liberdade evoluiu no transcorrer do tempo evolutivo. Nossa amígdala, por exemplo, um órgão que compõe o sistema cerebral e que tem a função de estabelecer o registro do medo, possivelmente

---

<sup>106</sup> Poderia ser descrito, por exemplo, como duas ou mais construções de narrativa que de algum modo (talvez por causa de um trauma específico) burlou o processo algorítmico que caracteriza o modelo pandemônico.

<sup>107</sup> Esse problema não é novo na filosofia, tendo, por exemplo, sido considerado por Locke.

<sup>108</sup> Uma implicação do *self* dennettiano é o problema da responsabilidade moral. Se realmente somos, enquanto *self*, uma ficção, uma edição de processos pandemônicos, uma construção inconsciente, como poderíamos nos responsabilizar por nossas ações morais?

não é livre em nenhum sentido significativo – já que é um “robô” –, ainda assim agrega a mente humana a incumbência de evitar os perigos e as ameaças externas. Assim, temos um argumento que nos permite entender como é possível a evolução de um determinismo natural para um estado de liberdade, cujo todo desfruta de uma liberdade que suas partes não degustam.

A despeito do ceticismo direcionado à imputabilidade metafísica, é conferido ao *self* o controle do organismo. Sem embargos, a autoconsciência emerge em sua verdadeira consistência mental e cerebral<sup>109</sup> por intermédio das estruturas internas na interação com o ambiente, mas isso não impede de realizar a atribuição de estados mentais a partir de uma interpretação heterofenomenológica, identificando-o como uma “unidade global do sistema a quem é atribuída a autoria de sua vida mental e controle de suas intenções” (cf. De Paulo, 2018, p. 54).

Em Dennett (1991a), o *self* é construído a partir de aprendizado numa perspectiva em terceira pessoa. Essa é uma estranha dependência, haja vista que o cérebro não está cômico das atividades subjacentes à construção do “eu”, mas contrariamente a um conteúdo interno impessoal, a partir do qual sua essência permite realizar a edificação do “eu”. Nesse contexto, estamos diante de um conhecimento oriundo da abordagem heterofenomenológica – em terceira pessoa – do sujeito sobre si mesmo. Dito de outro modo, se configura como um tipo de conhecimento advindo de fora para dentro e não o contrário. O modo de aquisição desse tipo de conhecimento é calcado na observação do nosso comportamento. Assim, criamos uma representação de nós mesmos a partir de uma ação executada que será tingida de narrativas sobre como nos relacionamos com o ambiente, fazendo emergir o centro de gravidade narrativo abstrato sobre essas narrativas. Portanto, claro está que o que possibilita a distinção entre o *self* e outras representações do nosso ambiente é o modo como agimos, ou seja, nossa autorrepresentação se resume à identificação do modo como agimos. Com efeito, as narrativas realizadas pelo sujeito são tão importantes quanto as realizadas por outras pessoas, haja vista que adquirimos “através de amigos, conhecidos, parentes ou até inimigos, narrativas sobre seu próprio *eu* ou situações no ambiente que dizem respeito ao seu *eu* que são incorporadas à sua autoconcepção” (De Paulo, 2018, p. 55 – grifo nosso).

Alguns poderiam tentar refutar o conceito de consciência de identidade como centro de gravidade narrativa, afirmando se tratar de uma ficção teórica. Estão equivocados, dirá Dennett

---

<sup>109</sup> Como vimos, de acordo com Dennett (1991b), o cérebro seria o centro de comando do organismo, mas não existe nenhum ponto centralizador no cérebro, nenhum teatro cartesiano.

(1991a), pois esse aspecto é na verdade o seu grande mérito. Dennett (1991a, p. 429 – nossa tradução) nos traz um panorama que facilita mais o entendimento de seu pensamento:

Este não é o problema com os centros de gravidade; é a glória deles. São ficções magníficas, ficções que qualquer um teria orgulho de ter criado. E os personagens fictícios da literatura são ainda mais maravilhosos. Pense em Ishmael, em Moby-Dick. “Me chame de Ismael” é a forma como o texto abre, e nós agradecemos. Não chamamos o texto de Ismael, e não chamamos Melville de Ismael. Quem ou o que chamamos de Ismael? Chamamos Ishmael de Ishmael, o maravilhoso personagem fictício que pode ser encontrado nas páginas de Moby-Dick. “Me chame de Dan”, você ouve de meus lábios, e você concorda, não chamando meus lábios de Dan, ou meu corpo de Dan, mas me chamando de Dan, a ficção teórica criada por ... bem, não por mim, mas pelo meu cérebro, atuando em concerto ao longo dos anos com meus pais, irmãos e amigos.

De acordo com Dennett (1969), o aspecto contra-intuitivo que permite estabelecer uma analogia do *self* como ficção não enfraquece a sua teoria, mas a fortalece, haja vista que podemos realizar uma representação do “eu” multiforme por um determinado período de tempo segundo o qual é construído um leque de narrativas que irá compor nosso “eu”, calcado em diferentes interpretações e atribuições. A representação do *self*, portanto, não constitui o próprio *self*, mas é o alicerce segundo o qual delineamos nossas narrativas sobre nós mesmos, sobre o qual se estabelece o centro de gravidade narrativo que forma a nossa consciência de identidade. No entanto, há um agravante em um ponto da teoria de Dennett concernente ao problema da consciência de identidade.

Quando é realizada a pergunta pelo *self*, Dennett (1991a) nos apresenta uma resposta calcada em um método epistemológico para uma questão ontológica. Além disso, a noção do “eu” como centro de gravidade narrativo parece ecoar o existencialismo heideggeriano, onde o homem é entendido como projeto que se constrói no mundo (Davis, 2020, p. 26-27). Como implicação, a abordagem epistemológica de Dennett recai em uma concepção ontológica da consciência, apresentando o mesmo problema do dualismo. Dado que o fundamento do centro de gravidade narrativo é formado por um banco de dados constituído por memes e a manifestação da identidade pessoal, autorrepresentação ou comportamento, se dá através da linguagem, a analogia com a tese haideggeriana que descreve a “linguagem como a morada do ser” parece inevitável. Heidegger (2015, p. 85-87 – grifo nosso) nos posiciona ainda mais:

A *essência* da presença está em sua existência. As características que se podem extrair deste ente não são, portanto, *propriedades* simplesmente dadas de um ente simplesmente dado que possui esta ou aquela *configuração*. As características constitutivas da presença são sempre modos possíveis de ser e somente isso[...] A presença se determina como ente sempre a partir de uma possibilidade que ela é e, de algum modo isso também significa que ela se compreende em seu ser. Este é o sentido formal da constituição existencial da presença.

Esse aspecto da teoria dennettiana não seria problemático, caso o filósofo introduzisse um aspecto ontológico na sua teoria. Mas isso não aparece em seu argumento sobre a

consciência de si. Assim, na discussão sobre o *self* dennettiano, estamos diante do aspecto mais ontológico de uma teoria da mente que pretende abordar o problema sob a égide de uma metodologia majoritariamente epistêmica. Ao que parece, a sombra do dualismo também atormenta o pensamento de Dennett.

## CONSIDERAÇÕES FINAIS

A teoria de Dennett consiste em desenvolver uma filosofia da mente que tenha como background os trabalhos desenvolvidos no campo das ciências empíricas, mas intenciona solucionar o problema da coisa-mente sem a necessidade de uma teoria da consciência e de uma teoria do conteúdo – embora preserve esta distinção afim de estabelecer a caracterização de sua abordagem.

Com efeito, Dennett (1991a) inicia sua análise do problema ontológico da consciência averiguando quais são as limitações cruciais da doutrina cartesiana e atacando o conceito de substância pensante de Descartes, sugerindo que somos um emaranhado de “vozes” sussurrando em nossa mente. Não há, segundo o filósofo, um administrador central em nosso cérebro que confere unidade a nossa consciência – não há uma localização específica onde a consciência se manifesta. Os estados mentais (a experiência consciente), nos parâmetros do dualismo, não existem enquanto entidades, ou seja, a teoria do “teatro cartesiano” é equivocada e inviável para as pretensões de uma abordagem científica do problema. Nesse sentido, a intencionalidade não compõe a natureza ou a essência da consciência. O filósofo dá uma resposta naturalista ao problema da mente. Sua concepção de consciência é bastante peculiar – a consciência não seria um pensamento de um pensamento, mas uma narrativa do que ocorre na mente. Assim, tentar “explicar a natureza da consciência é explicar como se formam conteúdos mentais, ou seja, a experiência consciente é imanente ao próprio pensamento” (Teixeira, 2008, p.21).

Nesse sentido, a mente é o que podemos chamar de *alma informacional* – formada por minúsculos “robôs”. Como vimos, o funcionalismo parte do pressuposto de que determinadas características definem, de modo essencial, os estados mentais. São elas: 1) os efeitos ambientais exercidos sobre o corpo; 2) as relações de estados mentais com outros estados mentais; 3) e a relação de estados mentais com o comportamento corporal.

Segundo Dennett (1999), a consciência se manifesta em uma estrutura sequenciada de *outputs* e *inputs* (entrada e saída de dados) no cérebro, formando uma cadeia seletiva de informações. Influenciado profundamente pelos estudos da inteligência artificial e com o objetivo de construir os fundamentos para sua psicologia, o filósofo assimila a mente como algo que pertence ao campo do virtual, ou seja, como um sistema intencional (de ficções úteis) e não como uma entidade que possui existência própria.

Com efeito, os estados mentais não passam de eventos físicos, e o que há de comum entre esses atributos mentalistas compartilhados pela espécie humana é explicado pelo

estabelecimento da “descrição de sistema preditivo que é neutro ao fisicalismo e, por essa razão, inteiramente compatível com o fisicalismo” (Dennett, 2006, p. 26). Assim, a mente é um conceito operacional por ser concebida como “uma reconstrução racional da observação das sequências de comportamentos de um organismo ou dispositivo” (Teixeira, 2016, p. 12) e essa postura configura o aspecto instrumentalista da teórica dennettiana. Como se vê, sua concepção de consciência vai além do convencional:

Temos pouco acesso aos processos pelos quais as palavras “nos ocorrem” para dizer, mesmo nos casos em que falamos deliberadamente, ensaiando nossos atos de fala silenciosamente antes de pronunciá-los. Os candidatos a algo para dizer simplesmente surgem de onde não sabemos. Ou nos pegamos já os dizendo, ou nos pegamos verificando-os, às vezes descartando-os, outras vezes editando-os um pouco e depois dizendo, mas mesmo esses passos intermediários ocasionais não nos dão mais dicas sobre como os fazemos. Nós apenas aceitamos ou descartamos esta e aquela palavra. Se tivermos razões para nossos julgamentos, eles raramente serão contemplados antes do ato, mas apenas retrospectivamente óbvios. [...] Portanto, realmente não temos uma visão privilegiada dos processos que ocorrem em nós para nos levar do pensamento à fala. Eles podem ser produzidos por um pandemônio, pelo que sabemos (Dennett, 1991a, p. 304).

O pensamento de Dennett (1991a) é uma resposta plausível ao problema das outras mentes. No entanto, ele é acusado de negar o caráter introspectivo da consciência – seus críticos o acusam de abandonar gradativamente a discussão do testemunho da subjetividade – o que causa, no mínimo, um desconforto a quem adota sua teoria. Mas quando levamos em conta as teorias que consideram a consciência como um dado intuitivo, a proposta de Chalmers por exemplo, fica evidente que o modo pelo qual resolvemos os problemas fáceis, não são suficientes para liquidar o problema difícil. Apesar de toda explicação satisfatória dos fenômenos que englobam os chamados problemas fáceis (reprodução, desenvolvimento, crescimento, metabolismo, automanutenção, autodefesa imunológica), a consciência permaneceria um mistério.

Há um problema, de acordo com Dennett (1997), na tese que tenta fazer um paralelo entre o físico e o fenomênico. Todo aquele que concebe a experiência consciente como um princípio fundamental, como algo similar à energia e à matéria, comete um erro crasso. Segundo o filósofo, as características que envolvem uma experiência não nos permitem tomá-la como uma propriedade funcional e, por esse motivo, não poderiam ser considerados princípios de uma crença. Caso houvesse essa possibilidade, afirma o autor, qualquer qualidade seria suficiente para fundamentar uma teoria.

Se a mente (os *qualia*) é a causa dos comportamentos então temos que concluir que ela seria algo distinto do cérebro, mas se, contrariamente, o cérebro é a causa dos *qualia* então a mente é seu produto. Mas Dennett procura demonstrar que os *qualia* não existem. Embora, se

a hipótese da inexistência das vivências introspectivas seja de difícil aceitação, a concepção contrária não é nada confortável – não há lógica em uma experiência de si cuja descrição é impossível. Sobre esse ponto do pensamento dennettiano, Teixeira (2008, p. 25-26) nos traz um panorama que facilita o entendimento dessa questão:

Dennett expõe o caso de dois provadores de café, Chase e Sanborn, cujo trabalho consiste em assegurar que o sabor do café que eles provam se mantenha o mesmo. Ambos afirmam terem perdido o prazer que antes experimentavam com o sabor do café que provavam. Chase acha que deixou de apreciar o café porque se tornou um apreciador mais sofisticado, embora o café que ele prove tenha mantido o sabor inalterado. Sanborn acha que deixou de apreciar o café não porque o gosto tenha se alterado, mas porque seus sensores gustativos se alteraram. Segundo Dennett nem Chase nem Sanborn podem saber o que aconteceu, tampouco exprimir uma diferença entre cada um dos casos. Não se saberá nunca se houve uma mudança nos *qualia* ou nas disposições reativas. Se há os *qualia*, estes se tornam inatingíveis, praticamente quiméricos. Se o sujeito não é capaz de perceber a diferença, diz Dennett, é porque ela não existe.

Nessa perspectiva, a fenomenologia só considera existente aquilo que é percebido e, nessa condição, podemos desconsiderar os *qualia* – é uma falácia considerar a possibilidade de sentimentos inefáveis. Mas podemos realmente realizar uma descrição, através da linguagem, dos nossos sentimentos, numa perspectiva heterofenomenológica? Considere, por exemplo, o sabor do maracujá. De que modo poderíamos descrever o sentimento de que uma coisa é azeda a partir da linguagem? Como falar para alguém que nunca tenha experimentado o sabor azedo que esse alimento é ‘menos azedo’ do que, por exemplo, o limão? Por outro lado, caso a tese que rejeita os *qualia* fosse verdadeira, dirá um crítico de Dennett<sup>110</sup>, não teríamos necessidade de experimentar, por exemplo, um caviar ou um whisky para realizar um julgamento de sua qualidade. A leitura do rótulo seria suficiente. Essa tese se aproxima do pensamento de Wittgenstein observado nos Livros Azul e Marrom<sup>111</sup>. No entanto, essa postura não é um contraexemplo ontológico. Ela se mostra muito mais como um contraexemplo epistemológico. Abster-se dos *qualia* não é fácil. Seria possível realizar uma referência interna sem o auxílio de uma referência externa? Para resolver essa questão, a noção de sistemas intencionais é de grande valia.

Em Dennett (2006), o termo intencionalidade é utilizado num sentido peculiar. A intencionalidade pressupõe o resultado de uma ação inconsciente antes do pensamento ou, melhor dizendo, é uma programação que permite uma racionalidade expressa, por exemplo, em um desejo (pelo homem) ou objetivo (por um computador). O comentário de Dennett (2006, p. 37) a esse respeito é esclarecedor:

---

<sup>110</sup> O filósofo Sidney Shoemaker é um exemplo (cf. Teixeira, 2008, p. 26).

<sup>111</sup> Os Livros Azul e Marrom (Blue and Brown Books) são dois conjuntos de notas elaboradas durante palestras realizadas por Wittgenstein no período de 1933 a 1935. Eles foram mimeografados separadamente.

Vê-se o computador como um sistema intencional. Prediz-se o comportamento, nesse caso, atribuindo ao sistema a posse de determinada informação, e pressupondo que ele é regido por determinados objetivos, e então elaborando a ação mais razoável e apropriada com base nessas atribuições e pressuposições.

Nessa perspectiva, não faz sentido afirmar que sistemas intencionais realmente possuem crenças e desejos, mas que, tão somente, podemos atribuir ações volitivas, ou seja, só é possível realizar explicações e previsões de comportamento a partir de atribuições de crenças e desejos às estruturas que apresentem postura de projeto ótimo (cf. Dennett, 2006, p. 38). Assim, claro está, os comportamentos de sistemas intencionais são baseados em análogos de crenças e desejos, pouco importando o nome que se queira dar a essa posse epistêmica. A consciência até aqui considerada significa nada mais que um projeto ótimo em relação a um objetivo, ou uma hierarquia otimamente balanceada de objetivos e um conjunto de imposições. O sistema intencional humano é subvencionado por um banco de dados formado por memes<sup>112</sup>.

Uma das colunas que sustenta o funcionalismo homuncular de Dennett é a teoria dos memes, mas é também um dos pontos mais vulneráveis de sua tese sobre a coisa-mente. A afirmação que os memes são parasitas de cérebros causa um desconforto – sobre tudo para aqueles que tentam imaginar a possibilidade de que algo que possui status ontológico podem ser abrigados por algo físico.

Além disso, a inexequibilidade de um exame observacional dessas máquinas virtuais afasta a possibilidade de uma analogia mais precisa entre memes e genes. Como teriam surgido os primeiros memes? Há quem os aproxime à psicologia junguiana, por apresentarem uma interessante similaridade – calcada na noção do meme como um tipo de arquétipo. Mas isso apenas reforça as críticas, dado que essa analogia é de cunho ontológico. Não sabemos como mensurar um meme, posto que é possível realizar medições em unidades de informação. Também somos completamente ignorantes quanto à possibilidade ou não de “modelos epidemiológicos [...] ser transpostos para a caracterização de um fenômeno psicossocial como a disseminação da informação” (Teixeira, 2008, p. 24).

A despeito desses problemas, podemos concordar com Dennett (1997), que cada unidade celular – que pode ser descrita como um agente simples com atribuições limitadas – é privada de uma consciência assim como um vírus. Seria possível que a constituição de uma pessoa consciente genuína fosse estabelecida por um agregado desses homúnculos ignorantes?

---

<sup>112</sup> Os casos que corroboram a tese dos memes não se resumem, por exemplo, às canções que nos flagramos repetindo sem entender o porquê ou palavras que nos deixam incomodados quando não sabemos o significado. O que causa perplexidade é que “teorias conspiratórias e abdução por seres de outro planeta também seriam memes” (Teixeira, 2008, p. 24).

A atual ciência padrão parece corroborar com essa tese. Nesse sentido, devemos considerar plausível a hipótese de que somos descendentes de robôs, mas isso não implicaria que nós mesmos sejamos robôs, dado que o fato de sermos descendentes diretos de répteis, por exemplo, não implica que somos répteis. Se partimos da ideia de que não há uma característica miraculosa em nós (aquele defendido pelos dualistas e os fisicalistas incautos), temos que admitir que somos constituídos por robôs – ou somos formados por “uma coleção de trilhões de máquinas macromoleculares” (Dennett, 1997, p. 22).

Os problemas com os memes se refletem também em críticas técnicas, haja vista que a noção de máquina virtual, que precede o modelo *pandemonium*, pode apresentar algumas limitações. Dada sua natureza de máquina virtual, ela não tem uma localização específica no cérebro – muito menos um fantasma que observa a sua atividade. Sua função é reunir os temas elaborados pelos homúnculos especialistas, de modo que seja possível uma narrativa coerente para publicação. No entanto, ao conceber essa máquina como um programa (*software*) do cérebro, a tese dennettiana se vê diante de alguns problemas.

De certo, em consonância com um ponto clássico em teoria da computação, a composição de um programa é estabelecida por algoritmos e estrutura de dados – caracterizados, por sua vez, por um tipo de linguagem simbólica. Ainda que estejamos diante de uma estrutura de computação paralela, é necessário um projeto que defina as operações das várias conexões dessa rede em paralelo. Tampouco podemos desconsiderar a existência de uma intenção que perpassa o projeto – os resultados são desconhecidos pelo sistema<sup>113</sup>. Não podemos dispensar essa entidade, haja vista que ela é a única que possui a prerrogativa de considerar o sistema como tal. Dito de outro modo, “o processamento paralelo é eficiente no sentido de que cada parte não precisa ter a visão do todo, mas apenas desempenhar a tempo sua função – como numa linha de montagem de carros” (Teixeira, 2008, p. 30). Nesse sentido, a intenção do administrador é promover a integração das partes, pois o papel que essas partes exercem com eficiência não depende do conhecimento das atividades das outras partes – na verdade apresentam profunda ignorância.

A intenção de Dennett (1991a) é excluir o significador central. Isso só é possível se as atribuições que devem ser executadas por uma máquina serial (a mente) forem também realizadas por uma máquina paralela (o cérebro), sem a necessidade de uma coordenação centralizada. Mas Dennett é pouco específico na descrição do que seriam as suas máquinas

---

<sup>113</sup> Caso contrário, a entidade ou programador, criador da rede, estaria integrado ao sistema.

(serial e paralela) no que tange a montagem de uma sobre a outra. De forma bastante eloquente, Teixeira (2008, p. 30 – grifo nosso) afirma:

Certamente, ela não pode ser uma máquina de estados discretos, como a máquina de Turing, pois seu *software* não é um conjunto de algoritmos e de estruturas de dados, mas sim um conjunto de memes dados pela linguagem e pela cultura. Ou seja, o termo "máquina virtual" é usado num sentido excessivamente amplo e vago que certamente não seria aceito pelos cientistas da computação.

A máquina joyceana não é equivalente à máquina de Turing. Nossas mentes apresentam dificuldades, se comparada com um computador, para abordar problemas que necessitam um modo serial de resolução (cf. Teixeira, 2008, p. 30).

Encerrando as críticas técnicas ao pensamento dennettiano, concernentes à montagem de sua máquina híbrida, nos resta verificar até que ponto sua doutrina – o seu modelo *pandemonium* – apresenta um grau de realismo.

A crítica da neurociência aos modelos computacionais têm sido cada vez mais frequente. Nos últimos anos, há uma defesa gradual de uma concepção que toma o cérebro como uma espécie de máquina eletroquímica. Não obstante, pesquisas recentes que investigam o aspecto neurofisiológico do neurônio tendem a corroborar o modelo *pandemonium* dennettiano. No modelo apresentado por MacCulloch e Pitts, por exemplo, um neurônio é similar às entradas lógicas dos chips. Portanto, existiria apenas “uma passagem para corrente elétrica, e assim sendo, a porta lógica só podia ficar em duas posições – *on* e *off*, ou 1 e 0, respectivamente” (Teixeira, 2008, p. 31 – grifo nosso). Atualmente, no entanto, os estudos apontam para uma estrutura muito mais complexa dos neurônios. Nessas células nervosas, há filamentos que exercem a atribuição de amplificadores não lineares de frequências e voltagens. Essa descoberta refuta a tese de que os dendritos são mecanismos passivos e que realizam simplesmente a somatória linear dos sinais de *input*. Esse aspecto funcional do cérebro ainda é um mistério – temos muito a desvendar sobre o que parece ser uma música inaudível, quase imperceptível. Em detrimento das portas lógicas, o que temos na verdade é um sem número de minúsculas entradas, que não apresentam uma estrutura funcional em *on* e *off*, mas uma passagem cujos grupos debatem sobre a necessidade da permanência de mantê-la ou não aberta. Esse *pandemonium* ocorre em cada neurônio, em um tempo estimado em milissegundos, que tem que estabelecer um vencedor interino.

Com efeito, o pensamento de Dennett se revela como uma parte importante das pesquisas sobre neurociência computacional, haja vista que o modelo *pandemonium* se mostrar como uma boa descrição do que ocorre no cérebro, por estar mais próximo da realidade cerebral (cf. Teixeira, 2008, p. 32).

De acordo com Dennett (1991b), a resposta para a questão de como o cérebro pode abrigar a consciência, muitas vezes tem sido tomada como uma pergunta retórica, levando à crença de que esse problema está além do entendimento humano. O filósofo pretende demolir essa postura cética. Ele nos explica que podemos entender como a exagerada quantidade de atividade cerebral corresponde à experiência consciente. A menos que o leitor aceite a concepção do cérebro como uma espécie de computador, essa explicação não será coerente com as pesquisas das várias interfaces que abordam a questão. Nesse sentido, as muletas de imaginação nos são fornecidas pelos conceitos da ciência da computação, que nos ajudam a tatear na sala escura que envolve nossa fenomenologia – nossa experiência subjetiva – e o conhecimento que temos sobre nosso cérebro (cf. Dennett, 1991a, p. 433 – nossa tradução).

Ao considerar nossos cérebros como um complexo sistema de processamento de dados, Dennett acredita que podemos nos afastar gradativamente do modo abstruso que até agora foi abordada a questão e traçar o caminho mais seguro para entender como o nosso cérebro produz todos os fenômenos mentais. Embora a abordagem escolhida pelo filósofo seja interessante, devemos estar sempre alerta sobre como lidar com o problema para evitar recair em erros antigos – como por exemplo, considerar um meio-termo central, os *qualia* como existente – e o próprio filósofo admite que há confusões residuais e erros no esboço que apresenta, mas salienta que com sua teoria estamos diante do melhor caminho. Dennett (1991a, p. 455 – grifo nosso) ressalta que:

Minha explicação da consciência está longe de ser completa. Pode-se até dizer que foi só um começo, mas é um começo, porque quebra o encanto do círculo encantado de ideias que fazia parecer impossível explicar a consciência. Eu não substituí uma teoria metafórica, o teatro Cartesiano, por uma teoria não metafórica ("literal, científica"). Tudo o que fiz, na verdade, foi substituir uma família de metáforas e imagens por outra, trocando no teatro, a testemunha, o meio central, a imagem, por *software*, máquinas virtuais, rascunhos múltiplos, um pandemônio de homúnculos. É apenas uma guerra de metáforas, você diz – mas metáforas não são “apenas” metáforas; metáforas são as ferramentas do pensamento. Ninguém pode pensar sobre a consciência sem elas, por isso é importante equipar-se com o melhor conjunto de ferramentas disponíveis. Veja o que construímos com nossas ferramentas. Você poderia ter imaginado sem elas?

Mas quando são realizadas analogias com sistemas computacionais, as críticas nos parecem pertinentes. O funcionalismo dennettiano parece flertar com uma concepção de consciência como uma leitura interna e/ou externa de um padrão de comportamento – a capacidade de identificar alguma coisa, ou seja, ter consciência de algo depende do reconhecimento desse algo. O que é ver, se não termos consciência do que estamos vendo? O que é degustar, se não termos consciência de que estamos saboreando alguma coisa? Nesse sentido, é preciso “baixar” o conhecimento necessário: chamamos isso de aprendizagem. Mas

um programa de computador já traz consigo um “passo-a-passo” para atingir seu propósito específico. Caso surja um novo problema – cujo programa em questão não foi projetado – ele trará. A inteligência consciente possui a prerrogativa de, ante a um problema não previsto, reelaborar sua estratégia (escamotear, procrastinar, desenvolver mecanismos de contenção, elaborar novos conceitos, etc.). Um robô é capaz de realizar uma representação para si mesmo? Um computador só atingirá esse nível operacional ao se apropriar de um programa que escreva programas. No entanto, o nível de reprodução artificial da mente ainda está no primeiro estágio dos três que, segundo pesquisadores de IA, são necessários. Sobre essa questão, Dennett (1997, p. 54) nos fornece mais detalhes:

Por que as mentes artificiais, como os corações artificiais, não poderiam tornar-se reais – concretizadas – a partir de quase qualquer coisa? Uma vez que percebemos o que as mentes fazem (o que as dores, crenças, e assim por diante, fazem), deveríamos ser capazes de produzir mentes (ou partes mentais) a partir de materiais alternativos que possuem essas propriedades. E tem parecido óbvio para muitos teóricos – inclusive eu – que o que as mentes fazem é processar informação; as mentes são os sistemas de controle dos corpos e para executar suas tarefas determinadas elas precisam coletar, discriminar, armazenar, transformar e processar a informação a respeito das tarefas de controle que executam.

De acordo com Dennett (2006), a mente é uma construção teórica útil, uma espécie de reconstrução racional fundamentada na observação das sequências de comportamento de um organismo ou dispositivo. Ele afirma que foi a psicologia popular que abriu o caminho para todo vocábulo mentalista, mas é na elaboração do conceito de sistemas intencionais que Dennett emprega a sua indústria. Tudo aquilo que for descrito como um sistema intencional deve ser portador de uma mente, ou seja, é um sistema que identifica – explica e predica – comportamentos com base em atribuições de estados mentais (desejos, crenças, medos, pressentimentos, intenções, etc.).

Dennett reúne suas ideias em torno de um sistema coeso e de grande valor para a análise contemporânea do tema. Fazendo emergir reflexões e conceitos que se agregam às pesquisas, mas que não esgotam os problemas inerentes ao exame da inteligência consciente.

## BIBLIOGRAFIA

- ABBAGNANO, N. **Dicionário de filosofia**. São Paulo: Martins Fontes, 2000.
- BUNNIN, N.; TSUI-JAMES, E. P. (Org.) **Compêndio de Filosofia**. Trad. de Luiz Paulo Rouanet. São Paulo: Loyola, pp.165-189, 2002.
- BROWN, J. R. **The laboratory of the mind: Thought Experiments in the natural sciences**. London: Routledge, 1991.
- BOUCHER, P. **Artificial intelligence: How does it work, why does it matter, and what can we do about it?** Scientific Foresight Unit (STOA), 2020.
- CESCON, E. **Considerações acerca do conceito de consciência nas ciências cognitivas**. Cognition-estudo: Rev. Elet. Filo. Vol. 10, nº 2, p. 160-73, 2013.
- CHALMERS, D. J. **The Conscious Mind: In Search of a Fundamental Theory**. New York: Oxford University Press, 1996.
- COOPER, R. **Thought Experiments**. *Metaphilosophy*, v. 36, n. 3, p. 328-347, abr. 2005.
- CHURCHLAND, P. M. **Matéria e Consciência: uma introdução contemporânea à filosofia da mente**. Trad. de Maria Clara Cescato. São Paulo: UNESP Ed., 2004.
- DENNETT, D.C. **Content and Consciousness**. London: Routledge & Kegan Paul Limited, 1969, pp. 90-96.
- \_\_\_\_\_. **Consciousness Explained**. Boston: Little, Brown and Company, 1991a.
- \_\_\_\_\_. **The Brain and its Boundaries**, (C. McGinn Review, The Problem of Consciousness), The Times Literary Supplement, 1991b.
- \_\_\_\_\_. **Review on The Rediscovery of The Mind**. In: The Journal of Philosophy 90(4), p. 193-205, 1993.
- \_\_\_\_\_. **A Perigosa Ideia de Darwin – A evolução e os significados da vida**. Trad. Talita M. Rodrigues. Rio de Janeiro: Rocco, 1995a.
- \_\_\_\_\_. **The Unimagined Preposterousness of Zombies: Commentary on Moody, O. Flanagan and T. Polger**, Journal of Consciousness Studies, vol. 2, no. 4, 1995b.
- \_\_\_\_\_. **Is Perception the 'Leading Edge' of Memory?** A. Spadafora (ed.) Iride: Luoghi della memoria e dell'oblio, anno. VIII, n. 14 Abr. 1995c, pp. 59-78.
- \_\_\_\_\_. **Tipos de Mentes: rumo a uma compreensão da consciência**. Trad. de Alexandre Tort. Rio de Janeiro – RJ, Editora Rocco LTDA, 1997.
- \_\_\_\_\_. **Memes: Myths, Misunderstandings and Misgivings**, Draft for Chapel Hill Lecture, October 1998.
- \_\_\_\_\_. **The Zombie Hunch: Extinction of an Intuition?** In: Royal Institute of Philosophy Millennial Lecture. 28 nov. 1999.

\_\_\_\_\_. **Brainstorms**: ensaios filosóficos sobre a mente e a psicologia. Trad. Luiz Henrique de A. Dutra. São Paulo: UNESP Ed., 2006.

\_\_\_\_\_. **Darwin's strange inversion of reasoning**. PNAS vol. 106, June 16th, 2009a, suppl. 1, pp. 10061-65.

\_\_\_\_\_. **Intentional Systems Theory**. Oxford Handbook of the Philosophy of Mind B. McLaughlin, A. Beckermann, S. Walter, eds., OUP, 2009b, pp. 339-50.

\_\_\_\_\_. **Intuition pumps and other tools for thinking**. W.W. Norton & Company. New York. 2012.

DE PAULO, G.V. **O Self como uma Construção Narrativa**: Identidade e Ilusão na Constituição do Eu. [Doutorado]. Universidade federal do Rio de Janeiro. Instituto de Filosofia e Ciências Sociais. Rio de Janeiro, 2018.

DESCARTES, R. **Meditações metafísicas**. Trad. Maria Ermantina Galvão, Martins Fontes. São Paulo, 2001.

DAVIS, B.W. **Heidegger**: conceitos fundamentais. Trad. Fábio Creder. Petrópolis, RJ. Vozes, 2020. p. 17-32.

EYSENCK, M. W. **Manual de psicologia cognitiva**. Trad. de Luís Fernando Marques Dorvillé, Sandra Maria Mallmann da Rosa. 7 ed. Porto Alegre: Artmed, 2017.

GARDNER, Howard. **A nova ciência da mente**. trad. Cláudia Malbergier Caon. São Paulo: Editora da Universidade de São Paulo, 1996.

GOMES, D.S. **Inteligência Artificial**: conceitos e aplicações. Rev. Olhar Cient. FAAr, Vol. 01. Nº 02. p. 234-246, 2010.

HEIDEGGER, Martin. **Ser e Tempo**. Trad. Márcia Sá Cavalcante. Ed. Vozes. Petrópolis, RJ.

LEAL-TOLEDO, Gustavo. **Neurônios-espelho e o Representacionalismo**. Rev. Filos., Aurora, Curitiba, v. 22, n. 30, p. 179-194, 2010.

\_\_\_\_\_. **Dennett and Chalmers**: argumentos e intuição. Trans/Form/Ação, (São Paulo), v.29 (2), 2006, p.123-132.

NAGEL, Thomas. **Como é ser um morcego?** Tradução Paulo Abrantes e Juliana Orione. Cadernos de História e Filosofia da Ciência, Campinas, série 3, v. 15, n. 1, p. 245-262, 2005.

NORTON, J. 1996. **Are Thought Experiments Just What You Thought?** Canadian Journal of Philosophy, 26:333–66.

PRATA, T. A.; LIMA FILHO, M. M. **Oscilações entre o reducionismo e o fisicalismo não-redutivo no naturalismo biológico de John Searle**. Trans/Form/Ação, Marília, v. 36, n. 2, p. 195-218, 2013.

PEREIRA, M. **Considerações sobre a epistemologia dos experimentos mentais**. Conjectura: Filos. Educ., Caxias do Sul, v. 20, n. 2, p. 181-197, set./dez. 2015.

RAMACHANDRAN, V. S. **O que o cérebro tem para contar**: desvendando o mistério da natureza humana. Trad. de Maria Luiza X. de A. Borges. Rio de Janeiro: Zahar, 2014.

REICHERT, Cláudio. **Identidade pessoal importa?** uma leitura de Paul Ricoeur e Derek Parfit. *Fundamento – Revista de Pesquisa em Filosofia*, n. 18, jan-jun – 2019.

SEARLE, J. **The Rediscovery of the Mind**. Cambridge Mass., London: MIT Press, 1992.

\_\_\_\_\_. **Intencionalidade**. São Paulo: Martins Fontes, 1995.

\_\_\_\_\_. **O mistério da consciência e discussões com Daniel C. Dennett e David J. Chalmers**, tradução de Andre Yuji Pinheiro Uema e Vladimir Safatle - São Paulo: Paz e Terra, 1998.

\_\_\_\_\_. **Mind**: a Brief Introduction. Oxford: Oxford University Press, 2004.

\_\_\_\_\_. **A Redescoberta da Mente**. Trad. de Eduardo Pereira Ferreira. São Paulo: Martins Fontes, 2006.

TEIXEIRA, J. F. **A Mente segundo Dennett**. São Paulo: Perspectiva, 2008.

TRIPICCHIO, A.; TRIPICCHIO, A. C. **Teorias da mente**. São Paulo. Tecmedd. 2004.