

**UNIVERSIDADE DE CAXIAS DO SUL  
CENTRO DE CIÊNCIAS HUMANAS E DA EDUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA**

**DANIEL PIRES NUNES**

**O LIVRE-ARBÍTRIO EM JOHN R. SEARLE: UMA  
CONTRAPOSIÇÃO DO NATURALISMO BIOLÓGICO AO  
FISICALISMO E AO FUNCIONALISMO**

Caxias do Sul

2014

**DANIEL PIRES NUNES**

**O LIVRE-ARBÍTRIO EM JOHN R. SEARLE: UMA  
CONTRAPOSIÇÃO DO NATURALISMO BIOLÓGICO AO  
FISICALISMO E AO FUNCIONALISMO**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Filosofia, da Universidade de Caxias do Sul, como parte dos requisitos para a obtenção do título de Mestre em Filosofia.

Orientador: Prof. Dr. Everaldo Cescon

Caxias do Sul

2014

Dados Internacionais de Catalogação na Publicação (CIP)  
Universidade de Caxias do Sul  
UCS - BICE - Processamento Técnico

N972L Nunes, Daniel Pires, 1975-  
O livre-arbítrio em John R. Searle : uma contraposição do naturalismo biológico ao fisicalismo e ao funcionalismo / Daniel Pires Nunes. – 2014.  
85 f. : il. ; 30 cm

Apresenta bibliografia.  
Dissertação (Mestrado) - Universidade de Caxias do Sul, Programa de Pós-Graduação em Filosofia, 2014.  
Orientador: Prof. Dr. Everaldo Cescon.

1. Ciência - Filosofia. 2. Filosofia da mente. 3. Ontologia. 4. Neurociências. 5. Searle, John R., 1932- . I. Título.

CDU 2.ed.: 001.1

Índice para o catálogo sistemático:

1. Ciência - Filosofia	001.1
2. Filosofia da mente	13
3. Ontologia	111.1
4. Neurociências	612.82
5. Searle, John R., 1932-	1SEARLE

Catalogação na fonte elaborada pela bibliotecária  
Roberta da Silva Freitas – CRB 10/1730



UNIVERSIDADE DE CAXIAS DO SUL

*"O livre-arbítrio em John R. Searle: uma contraposição do naturalismo biológico ao fisicalismo e ao funcionalismo."*

Daniel Pires Nunes

Dissertação de Mestrado submetida à Banca Examinadora designada pela Coordenação do Programa de Pós-Graduação em Filosofia da Universidade de Caxias do Sul, como parte dos requisitos necessários para a obtenção do título de Mestre em Filosofia. Linha de Pesquisa: Conceitos Fundamentais de Ética.

Caxias do Sul, 15 de dezembro de 2014.

Banca Examinadora:

Prof. Dr. Everaldo Cescon (orientador)  
Universidade de Caxias do Sul

Prof. Dr. Itamar Soares Veiga  
Universidade de Caxias do Sul

Prof<sup>a</sup>. Dr<sup>a</sup>. Cinara Nahra  
Universidade Federal do Rio Grande do Norte

CIDADE UNIVERSITÁRIA

Rua Francisco Getúlio Vargas, 1130 – B. Petrópolis – CEP 95070-560 – Caxias do Sul – RS – Brasil

Ou: Caixa Postal 1352 – CEP 95020-972 – Caxias do Sul – RS – Brasil

Telefone / Telefax (54) 3218 2100 – www.ucs.br

Entidade Mantenedora: Fundação Universidade de Caxias do Sul – CNPJ 88 648 761/0001-03 – CGCTE 029/0089530

## **DEDICATÓRIA**

Dedico este trabalho à minha esposa Rosário e aos meus filhos Dimítri e Thales, por terem suportado a minha ausência e a minha rabugice durante estes dois anos.

## **AGRADECIMENTOS**

Ao meu pai, Leonardo, cujo apoio foi determinante para que este trabalho fosse possível e à minha mãe, Miriam, pelo apoio incondicional.

À minha prima Michelle, ao Rafael, ao Davi e à tia Marília pela hospitalidade e o carinho nestes dois anos em Caxias do Sul.

Aos amigos Jerzy e Fabíola Brzozowski tanto pelas nossas conversas quanto pelas dicas e a ajuda nos momentos difíceis.

Aos amigos Lincoln, Dayana e Joanna pela ajuda, amizade e hospitalidade.

Aos professores Vanderlei de Oliveira Farias e Jerzy Brzozowski da UFFS que criaram e conduziram o grupo de estudos que plantou a semente do interesse por realizar este trabalho.

Ao Programa de Pós-Graduação em Filosofia, PPGFIL, da Universidade de Caxias do Sul pela oportunidade de realização de trabalhos em minha área de pesquisa e à CAPES pela provisão de auxílio com as taxas escolares para a realização do mestrado.

Aos colegas, às secretárias e aos professores do PPGFIL; aos primeiros pelo companheirismo durante o curso – sobretudo ao amigo Marcelo Lucas Cesco –, às seguintes pela atenção dispensada e aos últimos pelas instigantes discussões que nos permitiram crescer um pouco mais como seres humanos.

Ao meu orientador pela atenção e pelo incentivo.

Aos colegas do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul pelo apoio.

Enfim a todos que, de alguma forma, contribuíram para que este trabalho fosse possível, pois, com ele, um sonho foi realizado.

## RESUMO

Nesta dissertação, procura-se analisar se o naturalismo biológico de John Searle se apresenta como uma alternativa mais viável que as correntes fisicalistas e funcionalistas nas pesquisas que tratam da questão do livre-arbítrio. Sendo assim, toma-se, como estratégia de pesquisa, identificar os pressupostos destas linhas de pensamento da filosofia da mente, bem como as consequências filosóficas da adesão a cada uma delas. Para seguir tal caminho, toma-se como fio condutor o conceito de intencionalidade intrínseca. Entretanto, primeiramente define-se o que se entende por livre-arbítrio para então caracterizar de forma geral os posicionamentos fisicalistas e funcionalistas na filosofia da mente e tratar de como a questão do livre-arbítrio surge e pode ser crucial para tais correntes de pensamento. Posteriormente o naturalismo biológico é sintetizado (sobretudo no que tange à ontologia da consciência e à questão da intencionalidade) e contraposto ao fisicalismo e ao funcionalismo para, então, examinar a possibilidade do livre-arbítrio. Em tal contraposição, cada teoria é decomposta em seus enunciados para que os mesmos possam ser analisados criticamente. Nesta análise, verifica-se que o livre-arbítrio parece não encontrar espaço no cenário apresentado pelo fisicalismo e pelo funcionalismo. Defende-se que o naturalismo biológico searleano consegue esclarecer mais do que as outras duas correntes filosóficas como pode a ação livre ter a origem da sua motivação no que é externo ao estado mental que a faz ser realizada. A partir de tais constatações, avalia-se suas implicações éticas articulando as questões da intencionalidade intrínseca, do livre-arbítrio, da inteligência artificial forte e da responsabilidade moral a fim de concluir que às máquinas atuais não se pode atribuir responsabilidade moral por não serem capazes de intencionalidade intrínseca. Em seguida, argumenta-se pela origem evolutiva da intencionalidade e, por conseguinte, da moralidade. Neste sentido, defende-se também que a neurociência não elimina a responsabilização moral pois não prova que o livre-arbítrio é uma ilusão, ou seja, que tal ramo da ciência não contradiz o naturalismo biológico de John Searle.

**Palavras-chave: John Searle. Funcionalismo. Livre-arbítrio. Naturalismo biológico. Fisicalismo.**

## **ABSTRACT**

This dissertation aims to examine whether John Searle's biological naturalism is a more viable alternative to current physicalist and functionalist positions in dealing with the issue of free will. Thus, my strategy is to identify the assumptions of these lines of thought and their philosophical consequences. In order to accomplish this goal the concept of intrinsic intentionality is taken as a guide. I begin by defining what is meant by free will and go on to broadly characterize physicalist and functionalist positions in philosophy of mind. Then, I go on to show how the question of free will arises and can be crucial to such currents of thought. Subsequently, I summarize the biological naturalist position (especially regarding the ontology of consciousness and the question of intentionality) and oppose it to physicalism and functionalism in order to examine the possibility of free will. In this opposition, each theory is decomposed into its main tenets so that they can be critically analyzed. In this analysis, it appears that free will does not seem to find any room in the scenario presented by physicalism and functionalism. It is argued that Searlean biological naturalism is able to explain – better than the other two positions – how free action can be motivated by something that is external to the mental state which is itself performing the action. I then evaluate the ethical implications of these findings, articulating the issues of intrinsic intentionality, free will, strong artificial intelligence in order to conclude that current machines cannot be assigned moral responsibility, since they are not capable of intrinsic intentionality. Then, I argue for the evolutionary origin of intentionality and therefore morality. Finally, I argue that neuroscience does not eliminate moral responsibility since it does not prove that free will is an illusion, i.e., that this branch of science does not contradict John Searle's biological naturalism.

**Keywords: John Searle. Functionalism. Free will. Biological naturalism. Physicalism.**

## **LISTA DE FIGURAS**

Figura 1 – Razão externa operando na deliberação.....	64
Figura 2 – Situação racional ideal.....	64

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>9</b>
<b>2 O FISCALISMO E SEUS PROBLEMAS À LUZ DA QUESTÃO DO LIVRE- ARBÍTRIO.....</b>	<b>14</b>
2.1 O fiscalismo e suas variedades .....	14
2.2 Fiscalismo e livre-arbítrio.....	18
<b>3 O FUNCIONALISMO E SEUS PROBLEMAS À LUZ DA QUESTÃO DO LIVRE- ARBÍTRIO.....</b>	<b>23</b>
3.1 O funcionalismo e suas variedades.....	23
3.2 O funcionalismo computacionalista de Turing.....	25
3.3 O funcionalismo computacionalista e o livre-arbítrio.....	29
<b>4 A ABORDAGEM SEARLEANA DA QUESTÃO DO LIVRE-ARBÍTRIO.....</b>	<b>32</b>
4.1 O naturalismo biológico de Searle.....	32
4.2 A ontologia da consciência.....	34
4.3 Intencionalidade: a diferença entre a experiência de perceber e a de agir.....	38
4.4 Naturalismo biológico versus fiscalismo.....	47
4.5 Naturalismo biológico versus funcionalismo.....	54
<b>5 IMPLICAÇÕES ÉTICAS.....</b>	<b>63</b>
5.1 Livre-arbítrio, IA forte e responsabilidade moral.....	65
5.2 Neurociência e livre-arbítrio.....	67
5.3 Naturalismo biológico, moral evolucionista e universalidade das leis morais.....	71
5.4 A falácia naturalista.....	73
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>76</b>
<b>REFERÊNCIAS .....</b>	<b>82</b>

## 1 INTRODUÇÃO

A filosofia da mente é atualmente uma das áreas da filosofia mais férteis e instigantes. Entretanto, é errôneo achar que as questões com as quais ela se defronta são novas – que não se tratou de tais questões anteriormente. O fato é que as atuais tecnologias baseadas em novas descobertas da ciência deram novo vigor a determinados problemas. O problema mente-corpo, o da causalidade e o da ontologia da consciência, dentre outros, apesar de não serem novos, são discutidos filosoficamente mas recebem contribuições de outras áreas do conhecimento humano tais como as neurociências e a ciência da computação. Uma questão clássica que aparece como pano de fundo de todos esses problemas é a do livre-arbítrio. Pode ocorrer que toda uma corrente filosófica seja rejeitada por aparentemente nos levar a negar o livre-arbítrio.

Há diversas posições clássicas acerca da questão do livre-arbítrio. A primeira a ser abordada portanto será a que nega a sua possibilidade, a saber, a determinista. O determinismo parte do princípio de que tudo – estados de coisas ou eventos – tem uma causa. Desta forma a natureza, segundo esta concepção, possui leis causais que chamaremos doravante de leis da natureza. Dado um evento ou estado de coisas atual e considerando as referidas leis da natureza, é necessário que ocorra um determinado evento ou estado de coisas. Explicando de forma melhor, dado um evento A e um evento B e que as leis da natureza estabelecem uma relação causal entre os dois de tal forma que sempre que A ocorrer seguir-se-á a ocorrência de B; se A ocorre, então necessariamente B ocorrerá.

O determinismo não abre espaço para qualquer possibilidade de livre-arbítrio, visto que a vontade do agente nunca poderá ser diferente do que o estado de coisas anterior já determinou. Para tal posição, o estado de coisas atual (incluindo o estado cerebral ou o mental) já é determinado pelo estado de coisas anterior (que é único, no sentido de que não

poderia ser outro) e determinará um único estado de coisas posterior (único porque também não poderá ser outro). É uma cadeia causal que não abre possibilidade para a contingência: todo o encadeamento é uma relação de necessidade. John Searle explica a posição determinista da seguinte forma:

A tese do determinismo a respeito das ações consiste em que cada ação está determinada por condições prévias causalmente suficientes. Para toda ação, as condições causais da ação em tal contexto são suficientes para produzir tal ação. (SEARLE, 2001, p. 277; 2013, p. 297)

Mas qual é o problema que resulta do determinismo, visto que a visão científica de mundo parece estar de acordo com tal posição? O problema é que se todo o estado de coisas que antecede o nosso nascimento tem uma relação de necessidade com o estado de coisas posterior ao nosso nascimento (e aqui devemos incluir tanto os estados cerebrais como os mentais), não há a possibilidade da ação livre. Isto porque a ação livre pressupõe contingência; pressupõe tanto poder ser quanto poder não ser realizada. Assim, não havendo ação livre, não há como um agente ser responsável pelos seus atos. Se, portanto, o determinismo estiver correto, não há também espaço para a ética.

Uma segunda posição é a libertarista<sup>1</sup> que defende que apesar de haver leis da natureza, o determinismo causal não se aplica, em geral, às ações humanas. Desta maneira, a vontade do agente é geralmente autônoma, ou seja, livre de coerção ou constrangimento externo.

A tese da vontade livre, às vezes chamada de "libertarismo", enuncia que em algumas ações, pelo menos, as condições causais prévias da ação não são causalmente suficientes para produzir tal ação. Admitindo-se que ocorreu a ação, e que ocorreu por uma razão, de qualquer forma, o agente poderia ter feito algo diferente do que fez, dados os mesmos antecedentes causais da ação. (SEARLE, 2001, p. 277; 2013, p. 297)

Para o libertarismo, o determinismo causal aplicado às ações humanas é portanto falso ou, pelo menos, mal formulado.

---

<sup>1</sup> Sendo alguns representantes desta posição Robert Kane, Carl Ginet, Randolph Clark (cf. LOPES, 2009, p. 12).

Outra posição clássica perante a questão do livre-arbítrio é a do compatibilismo – da qual Daniel Dennett pode ser tido como um representante (DEL AMO, 2007) – segundo a qual a liberdade da vontade é perfeitamente compatível com o determinismo causal, ou seja, mesmo que haja leis da natureza que estabelecem um encadeamento causal entre eventos e estados de coisas, há como a vontade do agente ter um certo grau de liberdade em relação a isto. Para o compatibilismo haveria graus de constrangimento externo em relação ao agente: o livre-arbítrio seria, assim, uma questão de grau. No caso de o compatibilismo estar certo, há espaço para a ética. E outras questões clássicas como a da fraqueza da vontade tornam-se importantes.

Entretanto, para Searle (2001, p. 278), o libertarismo é incompatível com o determinismo.

A posição compatibilista<sup>2</sup> sustenta que [...] a liberdade da vontade é totalmente compatível com o determinismo. Dizer que uma ação está determinada é simplesmente dizer que possui causas como qualquer outro acontecimento, e dizer que é livre é simplesmente dizer que é determinada por certos tipos de causas e não por outros. (SEARLE, 2001, p. 277; 2013, p. 297).

Assim, para o autor (SEARLE, 2001, p. 278), o compatibilismo – pelo menos da forma como ele o concebe – deixa de lado ou não entende, por assim dizer, o problema do livre-arbítrio. Ou seja, para ele, o compatibilismo não é uma posição filosoficamente significativa.

Dadas as considerações feitas, o que entenderemos por livre-arbítrio doravante será o seguinte: o rompimento das relações causais que determinam que um evento A necessariamente seja seguido por um evento B conforme as leis da natureza. Portanto um agente livre é aquele que pode romper com as determinações causais, mas não de forma aleatória<sup>3</sup>.

---

<sup>2</sup> Conforme Costa (2000, p. 24 - 26), podemos dar como exemplos de compatibilistas Walter Stace e Harry Frankfurt.

<sup>3</sup> Dennett (2006, p. 388) defende que o acaso não existe. A aparente aleatoriedade em alguns eventos se deve ao fato de não conseguirmos identificar um padrão.

Neste trabalho, pretendemos comparar especificamente duas teorias – o fisicalismo e o funcionalismo (sobretudo o computacionalismo) – às teses de Searle quanto à possibilidade do livre-arbítrio para, a partir do resultado obtido, poder servir a novas pesquisas no campo da ética e demais áreas do conhecimento filosófico. Não faremos aqui portanto uma petição de princípio assumindo o livre-arbítrio como postulado para – a partir daí – analisarmos a consistência do naturalismo biológico. Neste diapasão pretendemos investigar ainda se há possibilidade para a existência da inteligência artificial forte e, assim, contribuir para as pesquisas científicas que hoje tomam como pressuposta a veracidade do funcionalismo computacionalista.

Mas por que comparar a teoria searleana justamente ao funcionalismo e ao fisicalismo? Isto se deve à observação (que há alguns anos motivou este trabalho) de que nos cursos de engenharia – sobretudo a elétrica e a da computação – é muito comum a opinião de que a solução de problemas passa pelo paradigma da inteligência e que, portanto, a mente humana deve ser uma espécie de computador que instancia um programa que seria, por sua vez, a mente. É muito comum também em tais cursos a opinião de que toda a realidade é física, ou seja, que as leis da Física poderiam dar conta da realidade a qual seria somente objetiva. O vocabulário até então aplicado somente a humanos ou a certas espécies de animais passou a ser comumente aplicado a máquinas – o que evidencia uma das seguintes alternativas: ou tal vocabulário não é adequado (não é literal, pelo menos) ou há o pressuposto de que máquinas possuem mentes no mesmo sentido que possuímos e, assim, fariam parte dos seres que poderiam ter livre-arbítrio e, por conseguinte, ser passíveis de responsabilidade moral.

Como o funcionalismo sustenta (resumidamente) que a mente é um programa de computador e como o fisicalismo defende que tudo é físico, tais teorias são as que mais convergem para a forma de pensar que é comum no meio das engenharias e da ciência da

computação. Sendo assim, elas serão contrapostas à de Searle – naturalismo biológico – acerca da possibilidade ou não do livre-arbítrio e, adicionalmente, da inteligência artificial forte, de modo que os argumentos de Searle possam ser melhor avaliados. Em tal contraposição, cada teoria será decomposta em seus enunciados para que os mesmos possam ser analisados criticamente.

Pretendemos caracterizar de forma geral os posicionamentos fisicalistas e funcionalistas na filosofia da mente e tratar de como a questão do livre-arbítrio surge e pode ser crucial para tais correntes de pensamento. Depois pretendemos mostrar a diferença entre a posição reducionista e a não-reducionista para, então, salientarmos suas virtudes e sobretudo seus problemas quanto à questão do livre-arbítrio. Então, mesmo que a questão ainda fique em aberto, verificaremos se o livre-arbítrio encontra espaço no cenário que as correntes do fisicalismo e do funcionalismo apresentam.

Logo após, para termos uma ideia geral da filosofia da mente searleana, faremos uma síntese do naturalismo biológico. Na sequência, analisaremos a teoria de Searle acerca da ontologia da consciência e os principais tópicos da sua teoria da intencionalidade, sobretudo quanto à diferença entre a experiência de perceber e a experiência de agir. Assim sendo, feita a análise do fisicalismo, do funcionalismo e do naturalismo biológico, faremos uma comparação crítica sobre tais teorias no que toca ao livre-arbítrio e suas implicações éticas. Ademais, verificaremos se a questão do livre-arbítrio tem implicações na questão da possibilidade ou da impossibilidade da existência da inteligência artificial forte – o que também poderá ter desdobramentos na área da ética.

Enfim, analisaremos se o naturalismo biológico de John Searle, comparado ao fisicalismo e ao funcionalismo, oferece uma maior possibilidade de compatibilização do livre-arbítrio com o princípio da causalidade<sup>4</sup> e indicaremos os principais avanços obtidos.

---

<sup>4</sup> Na natureza todo evento tem uma causa. E se um evento ou algo causa um determinado efeito, tal causa é suficiente para que tal efeito ocorra.

## **2 O FISCALISMO E SEUS PROBLEMAS À LUZ DA QUESTÃO DO LIVRE-ARBÍTRIO**

Eventualmente uma corrente filosófica pode ser rejeitada por aparentemente nos levar a negar o livre-arbítrio – este parece ser o caso do fiscalismo. Procuraremos portanto caracterizar de forma geral os posicionamentos fiscalistas e indicar como a questão do livre-arbítrio surge e pode ser crucial para tal corrente de pensamento. Tal caracterização servirá de base para a posterior análise comparativa com o naturalismo biológico searleano.

### **2.1 O FISCALISMO E SUAS VARIEDADES**

Diferentemente dos dualistas, aqueles que defendem uma posição fiscalista em filosofia da mente afirmam que a mente não é algo que esteja fora, por assim dizer, do mundo físico. Para eles os fenômenos mentais são físicos e uma teoria física completa, que desse conta de toda a realidade, poderia explicá-los. Segundo Maslin (2009, p.74), o fiscalismo “diz que os seres humanos são entidades inteiramente materiais, cujo funcionamento e propriedades podem ser completamente explicadas pelos conceitos e teorias extraídas de uma física idealmente completa”. O psicólogo e filósofo Diego Zilio (2010, p. 218) afirma que “o fiscalismo pretende, em poucas palavras, explicar a mente sem ter que ir além do mundo físico” e resume as posições fiscalistas da seguinte forma:

Em síntese, o fiscalismo se distingue pela tese de que tudo o que “existe” ou tudo o que é “real” no mundo espaço-temporal é um “fato físico” ou uma “entidade física” e de que as “propriedades” dos fatos físicos ou são propriedades físicas em si, ou são propriedades “constituídas/realizadas/compostas” por propriedades físicas. (ZILIO, 2010, p. 219)

É uma posição que não dá espaço a explicações dualistas. Rejeita qualquer discurso que tente conceber a mente como algo sobrenatural ou como uma substância separada do mundo físico. Contrapondo os fiscalistas aos dualistas, Maslin (2009, p.74) explica que para os primeiros “os seres humanos são inteiramente partes do mundo físico natural, não

fantasmas sobrenaturais inexplicavelmente ligados a máquinas corpóreas”. O que se segue desta posição é que se realmente os seres humanos são totalmente partes do mundo físico e se esse mundo físico somente é composto por eventos físicos, parece impossível que uma mente não física possa existir e ainda exercer influência no mundo físico. O problema seria aquele enfrentado por Descartes em seu dualismo de substâncias: se a mente for uma substância não física, como pode ter uma relação causal com o cérebro e, se tiver, para quê serve tal órgão então? Zilio (2010, p. 221) pondera que para os fisicalistas

não há espaço para qualquer tipo de evento não-físico como causa de eventos físicos. Se existirem eventos não-físicos, eles não fazem diferença no mundo físico, e se o fizerem é porque são, também, eventos físicos e é por conta dessa característica que eles possuem poder causal.

Então, por este caminho, basta tratar de eventos físicos para lidar com o mundo físico. Até porque, segundo o próprio Zilio (2010, p. 222), para os fisicalistas a “relação causal [...] deve ser buscada no mundo físico, especificamente nas ‘propriedades’ dos ‘eventos físicos’”. Assim, segundo tal concepção, as leis causais devem ser as da física. Mas esta é também uma posição compatível com o dualismo de propriedades, o qual defende que apenas há substâncias físicas, mas que estas podem ter propriedades físicas e não físicas. Nesta linha, a consciência – por exemplo – seria uma propriedade mental não física. O professor Tárík de Athayde Prata – filósofo brasileiro dedicado à filosofia da mente – diz o seguinte:

A ideia que dá a especificidade do fisicalismo não-reduutivo como uma forma atenuada de dualismo de propriedades é a ideia de que as propriedades mentais estão fortemente conectadas às propriedades físicas, de modo que aquelas não podem existir sem estas. Isso faz das propriedades mentais, apesar de sua irreducibilidade às propriedades que são físicas em sentido estrito (aquelas descritas pelas ciências naturais básicas), um tipo de propriedades “físicas” em sentido amplo (pois elas não poderiam ser instanciadas independentemente da instanciação de certas propriedades físicas). A relação entre as propriedades mentais e físicas segundo o fisicalismo não-reduutivo pode ser pensada de diversas maneiras, mas a mais difundida ao longo das últimas décadas tem sido a superveniência, um tipo de relação que Searle admite, em certo sentido, para explicar a conexão entre mente e corpo. (PRATA, 2013, p. 257)

Ou seja, para uma das formas de fisicalismo, a saber, o fisicalismo não-reducionista, “os eventos mentais ‘sobrevêm’ (isto é, as propriedades mentais não são idênticas a propriedades físicas) a eventos físicos, ao invés de eventos mentais serem “reduzidos” a eventos físicos” (CESCON, 2013, p. 81). Andrew Melnyk (2003, p. 69), resumidamente explica – de forma intuitiva, como ele mesmo diz – o conceito de superveniência na filosofia da seguinte forma:

o mental (por exemplo) sobrevém sobre o físico se e somente se, uma vez que os fatos mentais tiverem sido fixados, os fatos mentais são por conseguinte também fixados; o modo como as coisas são mentalmente não pode variar (e não simplesmente “não variam”) sem que haja variação também no modo como as coisas são fisicamente.

Então, um fenômeno sobrevém a um fenômeno subveniente quando a sua existência surge e depende deste último. Só pode haver mudança no fenômeno superveniente se houver uma mudança no subveniente. Entretanto o fenômeno subveniente – que serve de base – não pode ser alterado pelo superveniente. Esta parece ser a fonte de um dos problemas do fisicalismo na questão do livre-arbítrio – trataremos dele mais adiante. Antes, vejamos outra forma de fisicalismo: o reducionista.

A concepção reducionista defende que tanto a consciência quanto os pensamentos, a vida inconsciente e subconsciente são eventos cerebrais, os quais, por sua vez, podem ser reduzidos a eventos físicos. Nisto é que consiste a diferença entre um fisicalismo reducionista e um fisicalismo não-reducionista: para os primeiros tudo é físico ou pode ser reduzido ao físico (o que é uma forma de dizer que é físico também); já para os últimos, as propriedades não-físicas (como as mentais) sobrevêm às físicas. Mas para ambos não há a necessidade de se extrapolar o mundo físico para explicar a mente.

Como foi dito anteriormente, segundo o fisicalismo, as leis causais devem ser as leis da física e, pela simplicidade de tal posição, pode-se considerar que ela tem a Navalha de

Ockham<sup>5</sup> a seu favor – sobretudo a concepção reducionista, pois ela não considera que haja duas espécies de propriedades. Contudo, o fisicalismo reducionista provoca eliminações que são contraintuitivas pois, ao afirmar que somente existe o que é físico – inclusive que somente há propriedades físicas –, ou concebe a consciência e outros eventos mentais subjetivos como físicos ou os elimina. Eliminá-los é negar a existência do que nos é evidente: que somos conscientes. Paulo Abrantes (2005, p. 239), ao expor a posição de Thomas Nagel, afirma que “[o] fisicalismo, como é usualmente entendido, propõe reduções que não são aceitáveis porque eliminam justamente o que, na sua visão, é o mais característico do mental: a experiência fenomênica”. Por outro lado, conceber os eventos mentais subjetivos como físicos significa incluí-los em um quadro no qual a subjetividade estaria sujeita às leis causais da física. Ademais, colocado desta forma sintética, o fisicalismo não explicaria como eventos físicos poderiam ter intencionalidade. Ou seja, o fisicalismo não daria conta de explicar como os estados mentais, sendo estados físicos do cérebro, podem ser acerca do que é externo à mente. Assim como na questão da superveniência, na questão da intencionalidade parece que o fisicalismo é incompatível com posições plausíveis, tais como

a sugestão externalista amplamente endossada por certos filósofos da mente de que o conteúdo representacional das atitudes proposicionais é parcialmente constituído por suas relações a estados de coisas exteriores às cabeças de seus donos. (MELNYK, 2003, p. 70)

Assim, um dos problemas que tal posição enfrenta é o de explicar como os estados físicos do cérebro podem “ser sobre outros estados de coisas, inclusive aquelas que não existem” (MASLIN, 2009, p.87) – é o problema da intencionalidade.

Quanto à posição eliminativista, seria um tanto quanto estranho defender que os nossos estados mentais, a consciência e a subjetividade não existem. Acerca disso, John Searle (2006, p. 156) afirma que “Se uma pessoa não é consciente, não há como eu possa

---

<sup>5</sup> “*Entia non sunt multiplicanda praeter necessitatem*”. Frase atribuída a Guilherme de Ockham. (BIZARRO, 2006, p. 535).

demonstrar a consciência para ela; se é consciente, é muito mais inconcebível que ela pudesse duvidar seriamente de que fosse consciente”. Talvez pudéssemos afirmar que tanto os nossos estados mentais quanto a consciência e a subjetividade têm uma ontologia diferente dos estados ou eventos físicos, mas suponhamos que seja uma hipótese descartada a de que não existem, pois ela não parece ser razoável.

## 2.2 FISCALISMO E LIVRE-ARBÍTRIO

Reiterando, o fisicalismo reducionista ou concebe a consciência e outros eventos mentais subjetivos como físicos ou os elimina. Como a eliminação é uma hipótese rejeitada por motivos evidentes, o problema de a subjetividade estar sujeita às leis causais da física, bem como o da intencionalidade, fazem-se importantes. Na “carona” de tais problemas emerge o do livre-arbítrio. Isto porque, segundo o fisicalismo, as leis causais devem ser as leis da física e, se a subjetividade está sujeita a elas, como poderíamos afirmar que as nossas ações são livres se estão caracterizadas pela causalidade<sup>6</sup> concebida por aquela ciência?

Bertrand Russell (1976, p. 70) afirma que “[a] ciência moderna mostra que (sic) concepção tradicional de causa e efeito é fundamentalmente errônea, impondo-se ser substituída por uma noção diferente”. A concepção tradicional à qual Russell se refere é a de que se um evento A causa outro B, sempre que ocorre B poder-se-ia inferir um evento A que fosse temporalmente anterior a ele. Em tal concepção “uma causa [...] deve ser tal que em nenhuma circunstância concebível deixe de ser seguida pelo seu efeito” (RUSSELL, 1976, p. 70). Haveria uma relação de necessidade. Entretanto o livre-arbítrio pressupõe uma relação de contingência, isto é, utilizando termos aristotélicos, para haver livre-arbítrio deve-se ter a potência dos contrários<sup>7</sup>. Numa posição fisicalista poder-se-ia dizer que o cérebro causa a consciência, que ele causa os fatos mentais, ou melhor, que os fatos físicos do cérebro causam

<sup>6</sup> Pode-se perceber que as questões da ontologia da consciência, da causalidade e do livre-arbítrio estão intimamente relacionadas.

<sup>7</sup> Poder realizar e poder não realizar uma determinada ação. Tal noção é abordada na obra intitulada de *Ética a Nicômaco* (ARISTÓTELES, EN 1113b).

os fatos mentais – os quais também seriam físicos ou seriam propriedades dos fatos físicos. O fisicalismo – assim concebido – parece não poder escapar do determinismo pois, como Searle (2007, p. 23) define, “A tese do determinismo afirma que todas as ações são precedidas por condições causais suficientes que as determinam”.

Searle (2010, p. 6) afirma que “sabemos apenas que os processos cerebrais causam os estados de consciência [... e que...] o fato de não sabermos *como* isso acontece não significa que não saibamos *que* isso acontece”. Ele não se considera um fisicalista. Na verdade o filósofo estadunidense propõe um rompimento com as categorias clássicas, a saber, o próprio fisicalismo (tanto o reducionista quanto o não-reducionista), o dualismo, o monismo e o materialismo – dentre outras –, pois considera que elas “engessam” a análise dos eventos mentais. Para ele, o problema da ação da consciência sobre o corpo oferece muitas dificuldades para a sua solução por causa da “nossa adesão à herança cartesiana das categorias do mental e do físico” (SEARLE, 2007, p. 25). Ele defende que “junto com a tradição cartesiana, herdamos um vocabulário, e, com o vocabulário, um determinado conjunto de categorias, dentro das quais estamos historicamente condicionados a raciocinar sobre esses problemas”. (SEARLE, 2006, p. 25).

Mas, enfim, qual é o preço que o fisicalismo paga ao assumir as categorias cartesianas (com as quais Searle supõe não se comprometer)? Podemos dizer que tal concepção fica a dever, segundo Zilio (2010, p. 238), uma explicação do que significa ser “físico”. Não poderíamos aceitar uma resposta que defendesse que os fenômenos físicos são observáveis e que os não físicos não o são – até porque esta seria uma posição dualista e, na verdade, os fisicalistas defendem que só há o mundo físico. E, com efeito, os buracos negros não são observáveis, mas, mesmo assim, a Física teoriza sobre os mesmos. Enquanto não houver uma definição precisa do que é ser “físico”, haverá uma certa “nebulosidade” acerca do que é ser “não físico”. Mas talvez o maior preço que o fisicalismo paga surge da assunção de que os

eventos mentais, apesar de subjetivos, são físicos, pois, assim, deveria haver como estabelecer leis que os regem – o que aparentemente parece comprometer tal posição com o determinismo.

Vejamos o saldo parcial das dificuldades enfrentadas pelo fisicalismo. A primeira pode ser colocada assim: se os eventos mentais são físicos – havendo leis físicas que os regem – e se a concepção tradicional de causa e efeito for verdadeira, não há a possibilidade do livre-arbítrio. Então, ou o fisicalismo está correto e o livre-arbítrio não é possível, ou o livre-arbítrio é possível mas o fisicalismo falha pelo menos em algum aspecto, seja pela posição reducionista, pela pressuposição da superveniência<sup>8</sup> dos estados mentais em relação aos físicos ou até pela concepção tradicional de causa e efeito. Isto porque não poderia haver uma ação livre se estivesse estritamente submetida às leis físicas.

Quando dizemos que algo é “livre”, nosso significado não é preciso, a menos que possamos descrever de que é livre. Qualquer que seja ou quem quer que seja “livre” não está sujeito a uma certa coerção externa e, para ser exato, devemos explicitar esse tipo de coerção. (RUSSELL, 2010, p. 138)

Ou seja, um ser consciente somente poderia ser livre se as suas decisões, crenças, desejos – dentre outros estados intencionais – não estivessem submetidos ao determinismo das relações de causa e efeito da ciência da Física. Isto remete à outra dificuldade com a qual o fisicalismo se depara, a saber, a da intencionalidade. O fisicalismo não consegue explicar satisfatoriamente como os estados físicos do cérebro podem ser sobre outros estados de coisas. Mais ainda, alguns estados intencionais seguem regras impostas pela razão.

[...] dado o fato de que tais estados podem ser logicamente relacionados uns aos outros, estão sujeitos às restrições da racionalidade e da normatividade. Crenças, por exemplo, podem conflitar, e caso nos demos conta de um conflito entre crenças, então racionalmente devemos abandonar uma delas ou ambas. (MASLIN, 2009, p. 93)

---

<sup>8</sup> "Propriedades de um tipo F são supervenientes em relação às de outro tipo G quando as coisas são F em virtude de serem G". (BLACKBURN, 1997, p. 372).

O exemplo acima, dado por Maslin, mostra que determinadas relações não são físicas. Se alguém crê que ingerir ovos faz mal à sua saúde e age de forma a evitar tal alimento, este agir parte de um estado intencional. O conflito de crenças – como citado acima – não é uma relação física. A causalidade como é concebida pela Física não se aplica a tal relação.

Mas se a Física – como ciência – pretende “traduzir” a realidade com um conjunto de leis que seja o mais completo possível mas não puder dar conta da intencionalidade, talvez devamos concluir que os estados intencionais não seguem as leis da Física. Melhor dizendo, se um estado intencional  $I_1$  está estritamente relacionado a um estado físico  $F_1$  do cérebro mas  $I_1$  entra em conflito com um outro estado intencional  $I_2$  que, por sua vez, está estritamente relacionado a um estado físico  $F_2$  do cérebro; então o estado físico  $F_1$  está em conflito com o estado físico  $F_2$ ? Se sim, qual seria a natureza de tal conflito? Seguiria as leis causais descritas pela Física? Abrantes (2005, p. 237) expõe a posição de Chalmers acerca da relação mente-corpo dizendo que:

[...] existiriam leis psicofísicas, relações entre propriedades físicas e propriedades mentais, irreduzíveis às leis da física (pelo menos às atualmente conhecidas) [...]. Essas leis psicofísicas assegurariam, para Chalmers, o caráter naturalista da sua posição, por fazê-la compatível, desse modo, com o quadro de mundo que nos traça a ciência contemporânea. (ABRANTES, 2005, p. 237).

A tentativa de Chalmers não explica como o estado físico  $F_1$  se relacionaria com o estado físico  $F_2$ . Se são estados físicos, devem seguir as relações de causa e efeito que as leis da Física descrevem, independentemente da relação entre os estados intencionais  $I_1$  e  $I_2$ .

Em suma, o fisicalismo ou não consegue explicar o livre-arbítrio ou o nega, pois o problema da intencionalidade coloca-se – pelo menos por ora – insuperável para tal concepção. Assim, se o ato de deliberar acerca do que se deve fazer segue as regras da razão e é sobre algo que ainda não existe, ou seja, uma ação que ainda não se consumou, como pode depender apenas dos estados físicos (cerebrais)? Como podem ser os estados físicos

(cerebrais) suficientes numa cadeia causal para os estados físicos (cerebrais) imediatamente seguintes no tempo e tal concepção ser compatível com o livre-arbítrio? Parece que não o é. Sendo assim, o fisicalismo que aceita a superveniência esbarra nesta objeção: a do epifenomenalismo, ou seja, as ações deliberadas – baseadas na razão – seriam apenas ilusoriamente livres, pois dependeriam exclusivamente do estado físico do cérebro para a sua ocorrência. Isto porque o mental não teria poder causal sobre o físico – somente o contrário poderia ocorrer. Por outro lado, o fisicalismo que sustenta que ou tudo é ou pode ser reduzido ao físico, ou provoca eliminações que não podemos aceitar – pois implicariam afirmações absurdas como a de que não somos conscientes ou a de que não possuímos subjetividade – ou também é incompatível com o livre-arbítrio, pois se a subjetividade estiver sujeita exclusivamente às leis causais da Física, não há que se falar em indeterminismo nos eventos mentais. Portanto, mesmo que a questão ainda esteja em aberto, pelo fato de a ciência da Física não ser atualmente completa, o livre-arbítrio parece não encontrar espaço no cenário apresentado pelas correntes do fisicalismo.

### 3 O FUNCIONALISMO E SEUS PROBLEMAS À LUZ DA QUESTÃO DO LIVRE-ARBÍTRIO

Assim como procuramos anteriormente caracterizar de forma geral os posicionamentos fisicalistas relacionado-os com a questão do livre-arbítrio, pretendemos – no presente capítulo – fazer o mesmo para o funcionalismo. Para tanto é preciso primeiramente expôr a tese funcionalista da mente, mesmo que de forma sucinta, para então analisar os seus eventuais problemas acerca do livre-arbítrio – o que mais adiante será utilizado para avaliar as teses de John Searle sobre os mesmos temas.

#### 3.1 O FUNCIONALISMO E SUAS VARIEDADES

O funcionalismo em filosofia da mente – expresso de forma sintética – é a tese de que a mente é uma função. Robert Van Gulick (2011, p. 128) explica que “O funcionalismo em seu núcleo é a tese de que as mentes e os tipos mentais devem ser entendidos em termos dos papéis de funções que os estados e processos específicos desempenham dentro de sistemas devidamente organizados”.

Os funcionalistas defendem que a mente se caracteriza pelo que faz com os estímulos que recebe (*inputs*) acompanhados das respostas que fornece (*outputs*) em termos de comportamento ou de respostas corporais involuntárias tais como a produção de lágrimas. Entretanto, na concepção deles, para uma função ser implementada, não importa a base física, isto é, para o *software* ser implementado, não importa em que tipo de base física o *hardware* é organizado, mas sim se as saídas estão de acordo com o que a função deve fornecer conforme as entradas que lhe são inseridas. Então a ideia principal do funcionalismo é “a de que *qualquer* sistema físico poderia produzir vida mental desde que sua organização funcional o permitisse” (TEIXEIRA, 2008, p. 52).

Assim, há como classificar três formas de funcionalismo, a saber, o funcionalismo metafísico, o psicofuncionalismo e o funcionalismo computacionalista (ou funcionalismo estado-máquina). O primeiro defende que “o estado mental é especificado de maneira puramente formal [e que] a mente é uma caixa-preta, como os psicólogos a chamam, mediando *inputs* e *outputs*, mas opaca no que concerne ao que efetivamente ocorre em seu interior. Ela funciona bem, mas como ela funciona é algo desconhecido” (MASLIN, 2009, p. 131). O psicofuncionalismo defende que um estado psicológico ou mental equivale a um estado computacional da máquina de Turing<sup>9</sup>. Para tal posição “a Psicologia é uma ciência complexa” (CESCON, 2013, p. 76) e “deve empregar os mesmos tipos de explicações pragmáticas das ciências biológicas” (CESCON, 2013, p. 76) em que cada órgão do corpo humano tem a sua função. Já o funcionalismo computacionalista – que abordaremos com maior profundidade mais adiante – trata a mente humana como uma função implementada pelo cérebro e que processa informações, ou seja, a mente é um programa de computador sendo implementado no cérebro que faz o papel de *hardware*. Teixeira (1998, p. 48-49) explica o funcionalismo como é defendido pelos teóricos da inteligência artificial da seguinte forma:

O funcionalismo, enquanto tese geral defendida pelos teóricos da Inteligência Artificial, sustenta que estados mentais são definidos e caracterizados pelo *papel funcional* que eles ocupam no caminho entre o *input* e o *output* de um organismo ou sistema. Este papel funcional caracteriza-se seja pela interação de um estado mental com outros que estejam presentes no organismo ou sistema, seja pela interação com a produção de determinados comportamentos. O funcionalismo consiste, assim, num nível de descrição no qual é possível *abster-se* ou *suspender-se* considerações acerca da natureza última do mental, isto é, se ele é ou não, em última análise, redutível a uma estrutura física específica. A descrição das funções é uma descrição abstrata, que tem o mesmo estatuto da descrição de um *software* ou fluxograma que estipula quais as instruções que um computador deve seguir para realizar uma determinada tarefa.

---

<sup>9</sup> Pode-se ver mais sobre a máquina de Turing em Maslin (2009, p. 137-141).

Como dissemos anteriormente, as atuais tecnologias baseadas em novas descobertas da ciência deram novo vigor a determinados problemas, mas também popularizaram algumas crenças, como a de que a mente é uma espécie de programa de computador – tese do funcionalismo computacionalista que também é muito difundida nos ramos da ciência da computação e da engenharia. Tal tese muitas vezes é simplesmente tomada como verdadeira e serve como ponto de partida – por profissionais das áreas citadas – para eventuais pesquisas de novas tecnologias. Por isso nos focaremos nesta concepção e analisaremos seus pontos fortes e fracos, sobretudo quanto à questão do livre-arbítrio.

### **3.2 O FUNCIONALISMO COMPUTACIONALISTA DE TURING**

Alan Turing propôs a seguinte questão em seu artigo intitulado “*Computing Machinery and Intelligence*” (TURING, 1950, p. 433): “Podem as máquinas pensar?”. E para evitar respostas baseadas em opiniões majoritárias acerca dos significados que as pessoas dão às palavras “máquina” e “pensar”, ele propôs uma nova formulação baseada em um “jogo da imitação” que é jogado por um homem (A), uma mulher (B) e um interrogador (C) que fica em um quarto separado dos primeiros. O interrogador (C) deve tentar descobrir qual dos outros dois é o homem ou a mulher, e estes tentarão despistá-lo. Eles somente se comunicam por escrito. Mas o que aconteceria se o homem (A) fosse substituído por uma máquina? Será que o interrogador teria a mesma dificuldade em identificar o humano e a máquina em comparação com a situação em que tinha que descobrir quem era o homem e quem era a mulher?

Reformulada, a questão inicial passaria a ser: “Existem computadores digitais imagináveis que tivessem bom desempenho no jogo da imitação?”. Para Turing, se o interrogador não puder diferenciar a máquina do humano ao final do “jogo da imitação”,

então pode-se concluir que a máquina pode pensar. Como não existem perguntas neutras ou alienadas de uma posição a respeito do que significa pensar, em princípio as perguntas poderiam ser de qualquer natureza (Turing não deixa isto claro), conforme a resposta que se precisasse para a decisão sobre se o interrogado é um humano ou uma máquina.

Então ele analisou algumas objeções à sua opinião, sendo a primeira a teológica (TURING, 1950, p. 443-444) a qual afirma que pensar é uma função da alma humana imortal. Animais ou máquinas não poderiam portanto pensar pois não lhes teria sido dada uma alma imortal por Deus. A segunda (TURING, 1950, p. 444) é a que ele chama de “cabeças na areia”, ou seja, a de que seria terrível se as máquinas pudessem pensar e que tomara que elas não possam fazê-lo. Tais argumentos pareceram fúteis a Turing, que os descartou sem maiores preocupações. Já a objeção de que há certas perguntas para as quais as máquinas não terão respostas recebeu de Turing o que segue: nós mesmos damos respostas erradas para várias perguntas (TURING, 1950, p. 444-445). O que realmente não implica o fato de máquinas não poderem responder a certas perguntas. Mas será que as máquinas poderiam responder – em sentido estrito – a tais perguntas? Searle dirá que não no seu famoso argumento do quarto chinês (SEARLE, 2010, p. 16-18).

O argumento da consciência (TURING, 1950, p. 445-447), que só pode admitir que uma máquina possa pensar se a mesma puder expressar pensamentos ou emoções sentidas, foi classificado por Turing como solipsista. Já a objeção das várias incapacidades (TURING, 1950, p. 447-450), que defende que uma máquina tudo pode fazer, menos X (podendo esse X ser: apaixonar-se, cometer erros, ter senso de humor, ser bondoso ou amigável, aprender com a experiência, fazer algo novo, etc.) obteve a resposta de que esta ideia que as pessoas têm do que as máquinas podem ou não fazer se funda no princípio da indução científica, que é falível, ou seja, é uma generalização a partir daquelas que já foram observadas.

Para a objeção de que as máquinas não podem fazer algo novo ou então que não podem nos surpreender, Turing (1950, p. 450-451) pôs em dúvida, no primeiro caso, se podemos ter certeza que o que fazemos tem caráter de originalidade, considerando que pode ser fruto de “uma semente plantada pelo ensino”. Já no segundo caso, identificou uma falácia, pois o surpreendido é o observador (e a surpresa então está nele) e não o observado. A continuidade do sistema nervoso comparada aos sistemas discretos das máquinas digitais também foi usada como objeção e Turing (1950, p. 451-452) argumentou que no jogo da imitação a máquina responderá probabilisticamente e assim não dará margem para que o interrogador possa identificar tal diferença.

O argumento da informalidade do comportamento afirma que não há como criar um conjunto de regras capaz de abarcar todas as situações possíveis para governar o comportamento de um homem. Segue-se que o homem não pode ser máquina. Turing (1950, p. 452-453) concordou com a primeira afirmação, mas diferenciou “regras de conduta” de “leis de comportamento” afirmando ainda que não temos certeza da ausência das últimas.

A última objeção analisada naquele artigo é a da percepção extrassensorial, levada muito a sério pelo autor, pois considerou que há evidências estatísticas do fenômeno. Assim, afirmou que se a percepção extrassensorial for comprovada “a ideia de que nossos corpos se movem tão somente de acordo com leis conhecidas da física, juntamente com outras ainda não descobertas, mas de alguma forma semelhantes, seria uma das primeiras a desaparecer” (TURING, 1950, p. 453). O autor inglês considerou que poderia haver ligação telepática entre o interrogador e o computador e, para que o jogo da imitação não fosse burlado, seria necessário que as salas dos participantes fossem à prova de “poderes” extrassensoriais.

Essas foram as objeções às quais Turing procurou refutar. Entretanto ele reconheceu (TURING, 1950, p. 454) que seus argumentos não são de todo convincentes para defender seus pontos de vista e que, por isso, seus esforços foram mais no sentido de apontar as

falácias das opiniões contrárias às suas. Contudo ele reforçou que o problema de se criar máquinas inteligentes e que, como os humanos, também aprendam é sobretudo de programação. Assim, o funcionalismo computacionalista de Turing concorda com a tese de que computadores programados de forma adequada têm mentes no mesmo sentido que os humanos têm, pois, segundo ele, os dois instanciam funções<sup>10</sup>. Andrew Hodges (2001, p. 42-43) sintetiza o que Turing defende da seguinte forma:

A tese de Turing é que as únicas características do cérebro relevantes para o pensamento ou a inteligência são aquelas situadas no nível de descrição da máquina-de-estado-discreto. A materialização física particular é irrelevante. [E ainda] toda característica do cérebro relevante para o pensamento pode ser captada numa tabela de comportamento e, assim, simulada por um computador.

No mesmo sentido Margoni (2013, p. 58) afirma que

O resultado dessa grande abstração sobre a qual se apoia o modelo computacional da mente foi abordar a cognição e a consciência como se essas pudessem ser produzidas por máquinas idealizadas, desde que essas fossem funcionalmente equivalentes a um sujeito humano.

Ou seja, poderíamos concluir, conforme tal afirmação, que se uma máquina se comporta como um ser humano, então ela possui consciência e cognição da mesma forma que tal ser. Para tal modelo (computacional), haveria uma relação estreita entre o modo como as máquinas funcionam – com *hardware* e *software* – e como os humanos “funcionam”. Entretanto, se assim é, pode ser que a ontologia da mente humana também tenha a ver com programas de computador. Então teríamos uma grande questão a ser resolvida que é a do livre-arbítrio humano. Searle (2010) faz duras críticas a esta concepção computacionalista da mente – trataremos delas com maior profundidade adiante.

---

<sup>10</sup> Instanciar uma função aqui é tomado no sentido de se concretizar ou materializar a execução de uma determinada função a partir de uma determinada configuração física. Uma função pode ser executada por mais de uma “base” física.

### 3.3 O FUNCIONALISMO COMPUTACIONALISTA E O LIVRE-ARBÍTRIO

Como foi exposto anteriormente, Turing (1950) diferenciou “regras de conduta” de “leis de comportamento”. Por “regras de conduta” ele entende as que regem a nossa ação e das quais temos consciência como, por exemplo, “levantar a mão para pedir permissão para falar em sala de aula”. Por sua vez, as “leis de comportamento” são as leis causais da natureza aplicadas ao corpo humano – mais especificamente ao cérebro –, acerca das quais Turing pensa que temos maior resistência a nos convencermos de que não somos governados. Para Turing (1950) é sobretudo um problema de programação o de se criar máquinas inteligentes. Mas programas de computador somente estabeleceriam “leis de comportamento”: isso significa que todas as possíveis “regras de conduta” teriam estatuto de “leis de comportamento”. Podemos considerar a partir daí, como hipóteses, as seguintes: ou computadores programados de forma adequada não têm mentes no mesmo sentido que os humanos as têm ou não há espaço para o livre-arbítrio. Por quê? Porque se a primeira hipótese estiver correta, o problema do livre-arbítrio humano nada tem a ver com programação e as máquinas não poderiam ter esta capacidade. Por outro lado, pela segunda hipótese, se o estatuto ontológico da mente tiver relação com programação, as “regras de conduta” poderiam ser reduzidas a “leis de comportamento”.

De certa forma, o problema parece ser o mesmo do fisicalismo – ou pelo menos muito parecido –, ou seja,

se todo o (sic) sistema físico segue as leis de Newton, se todo sistema físico é uma máquina, se o corpo humano é um sistema físico que segue as leis de Newton, logo, o corpo humano é uma máquina. (MARGONI, 2013, p. 40)

Com efeito, segundo Tripicchio (2004, p. 86), “o funcionalismo implica uma postura monista materialista não-reducionista” porque somente admite a existência do que é físico. Ou melhor, para o funcionalismo os estados funcionais não são propriamente estados físicos

mas são corporificados, instanciados, em estados físicos (MASLIN, 2009, p. 147). Portanto o funcionalismo pode ser visto como uma modalidade, por assim dizer, do fisicalismo. E desta maneira traz consigo as mesmas dificuldades acerca do livre-arbítrio enfrentadas por aquela concepção.

Então podemos identificar sérios obstáculos que o funcionalismo enfrenta sobretudo quanto à questão do livre-arbítrio. Começemos admitindo (como descrito anteriormente) que o corpo humano é uma máquina e que somente existe o que é físico, assim, forçosamente teríamos de concordar com o que explica Cescon (2013, p. 7), ou seja, que “De acordo com o funcionalismo estado-máquina, a natureza do estado mental é autômata”.

Vejamos. Pode-se entender “estatuto ontológico” como o “modo de ser” de alguma coisa. Mas o modo de ser das funções é o mesmo do dos objetos físicos? Ou então dos objetos mentais? Segundo Jaegwon Kim (2011), não. Para ele, podemos extrair as seguintes teses do funcionalismo: (I) Propriedades mentais são propriedades funcionais; (II) Uma função não é nem mental nem física. Considerando então que o funcionalismo afirma que computadores programados de forma adequada têm mentes no mesmo sentido que os humanos têm, pois os dois instanciarão funções. Levando em conta também que, para os funcionalistas, propriedades mentais são propriedades funcionais (MASLIN, 2009) e que uma função não é nem mental nem física, poderíamos dizer que “o funcionalismo, ao identificar os estados mentais com papéis funcionais, deixa de fora os *qualia*<sup>11</sup>” (COSTA, 2005, p. 30), ou seja, a subjetividade dos estados mentais – o que pode ser um problema quanto ao livre-arbítrio. Haveria como compatibilizar livre-arbítrio e ausência de subjetividade?

Além disso, a questão de os estados mentais poderem ser sobre outros estados de coisas – questão da intencionalidade –, bem como a dos *qualia* invertidos (PUTNAM, 2002, p. 267-277) parecem ser, por enquanto, barreiras intransponíveis ao funcionalismo

---

<sup>11</sup> Os *qualia* são estados qualitativos, subjetivos e internos, ou seja, conforme o próprio Searle (2010, p. 33-34) o problema dos *qualia* e o da consciência são idênticos. Isto porque, segundo ele, não existe nada na consciência além dos próprios *qualia*.

(dificuldades comungadas com o fisicalismo). O problema dos *qualia* ausentes e da intencionalidade foram especialmente abordados por Searle e suas críticas serão analisadas nos próximos capítulos.

#### **4 A ABORDAGEM SEARLEANA DA QUESTÃO DO LIVRE-ARBÍTRIO**

Para que se possa analisar se o naturalismo biológico de John Searle – contraposto ao fisicalismo e ao funcionalismo – oferece uma maior possibilidade de compatibilização do livre-arbítrio com o princípio da causalidade, faz-se necessário primeiramente entender, em linhas gerais, o que caracteriza o naturalismo biológico de Searle e como tal autor concebe ontologicamente a consciência. Isto porque, quando a questão do livre-arbítrio é tratada, necessariamente surge a questão de como podemos causar conscientemente um movimento corporal, por exemplo. Por trás desta questão está outra, a saber, a de como estados conscientes podem causar mudanças no que consideramos “mundo físico”. Se tivermos esta questão esclarecida, teremos maior facilidade para tratar do quão livres são os estados conscientes – e se são suficientes – para, por exemplo, realizar um movimento corporal.

##### **4.1 O NATURALISMO BIOLÓGICO DE SEARLE**

Naturalismo biológico é o nome dado por John Rogers Searle à sua posição filosófica que defende que “nem toda a realidade é objetiva; [que] parte dela é subjetiva” (SEARLE, 2006, p. 32). O filósofo estadunidense deu um nome à sua teoria acerca do problema mente-corpo para diferenciá-la das posições filosóficas tradicionais que, segundo ele, têm como pressupostas as categorias cartesianas. Segundo Maslin (2009, p. 165), o que Searle pretende ao utilizar o termo “naturalismo” para nomear a sua teoria é ressaltar que “eventos mentais fazem parte integral do mundo natural, e não estão, em sentido algum, além ou fora dele”. Para Searle os eventos mentais têm sua ontologia explicável sem fazer referência a eventos que estejam fora do mundo natural – o que elimina o dualismo cartesiano e não necessita de qualquer referência a alguma divindade para salvar a sua teoria. Ele rejeita tanto o dualismo de propriedades quanto o de substâncias, assim como o materialismo e o monismo pois

considera que tais concepções estão baseadas nas categorias cartesianas (SEARLE, 2006, p. 45) – que são, segundo ele, a origem dos equívocos.

Searle (2006, p. 83) considera a consciência como uma irreduzível<sup>12</sup> e avançada capacidade – dentre outras capacidades físicas – no nível macro que emerge de alterações no nível micro do cérebro. Afirma que, de certa forma, a consciência é ao mesmo tempo causa e efeito relativamente aos eventos que ocorrem no cérebro e tem uma característica importante: a subjetividade. Isto porque os fenômenos mentais são características de nível macro do que acontece em nível micro no cérebro, ou seja, “a mente e o corpo interagem, mas não são duas coisas diferentes” (SEARLE, 1984, p. 33). Tal afirmação pode ser um tanto quanto problemática, pois, se mente e corpo interagem, como podem não ser duas coisas diferentes? Como Searle poderia explicar tal interação não admitindo alguma forma de dualismo? Tárík de Athayde Prata o acusa de incorrer em uma outra forma de dualismo, pois segundo ele, Searle considera

a subjetividade como uma característica que distingue os fenômenos mentais conscientes de tudo mais no universo [...], o que, a despeito de sua afirmação de que o mental é físico, resulta em uma forma de dualismo, entre subjetividade e objetividade ontológicas. (PRATA, 2011, p. 572)

Somos forçados a concordar com tal afirmação, pois, conforme o naturalismo biológico (SEARLE, 2006, p. 34), podemos fazer o seguinte esquema:

- (1) “Fenômenos mentais estão essencialmente relacionados à consciência”.
- (2) “a consciência é essencialmente subjetiva”.
- (3) Então “a ontologia do mental é essencialmente uma ontologia de primeira pessoa”.

Explicando melhor, os fenômenos mentais conscientes (1) têm a característica de serem subjetivos (2), o que significa que eles só têm existência quando vivenciados pelo ser consciente, ou seja, o seu modo de ser é de primeira pessoa (3). Assim, podemos dizer que “a asserção de que toda a realidade é objetiva é, neurobiologicamente falando, simplesmente

---

<sup>12</sup> Conforme os nossos modelos de redução, pois os mesmos procuram eliminar justamente o que é subjetivo. (SEARLE, 2006, p. 176-179).

falsa” (SEARLE, 2006, p. 32). Ou seja, se a investigação científica procura eliminar o que é subjetivo e se parte da realidade é subjetiva, a ciência, como a concebemos, não dá e não dará conta de toda a realidade. Isto implica também na incapacidade de tal ciência abordar o mental.

Resumindo, o naturalismo biológico é a concepção de que parte da realidade é subjetiva e de que portanto a subjetividade é um fenômeno natural. É uma concepção que procura não eliminar a subjetividade, a consciência, a intencionalidade nem a causação mental<sup>13</sup>, pois, conforme o filósofo de Berkeley (na Califórnia), qualquer teoria que deixa uma dessas características de lado deve ter cometido algum erro (SEARLE, 1984, p. 22-23). Entretanto, será que o naturalismo biológico de John Searle, comparado ao fisicalismo e ao funcionalismo, oferece uma maior possibilidade de compatibilização do livre-arbítrio com o princípio da causalidade? Esta é a questão à qual nos dedicaremos na presente pesquisa.

#### **4.2 A ONTOLOGIA DA CONSCIÊNCIA**

A consciência como objeto de estudo, por não ser acessível do ponto de vista de terceira pessoa, parece escapar às investigações de qualquer ciência que conceba a realidade como sendo somente física. Uma característica fundamental da posição searleana (SEARLE, 1998, p. 24) acerca da ontologia da consciência é a de que ela é irreduzível a processos físicos, apesar de ser física<sup>14</sup>. As tentativas de eliminar a consciência reduzindo-a a alguma outra coisa, como a um programa de computador (funcionalismo computacionalista) ou a estados mentais que são descritos por termos puramente “físicos”, seriam para Searle (1998, p. 25) malsucedidas. Isto porque intrinsecamente negariam que temos estados internos qualitativos e subjetivos. A consciência então, por ser irreduzivelmente subjetiva, não é tida como um objeto

---

<sup>13</sup> Apesar de considerar a relação entre o micro e o macro como muito próxima para ser causal da forma como a ciência concebe a relação de causa e efeito.

<sup>14</sup> Física no sentido de fazer parte do mundo físico.

adequado de estudo por uma ciência que elimina a subjetividade – que procura ser o mais objetiva possível. Ou seja, a ciência procura (idealmente) eliminar da explicação de um determinado fato a subjetividade de quem explica, tornando-a assim (a explicação) o mais objetiva possível. Portanto a nossa herança cartesiana nos leva a estudar a mente descartando toda a sua subjetividade, eliminando a consciência. Há aqui duas posturas possíveis de serem assumidas, conforme tal herança: ou a consciência deve ser eliminada como realmente existente ou deve ser eliminada como objeto de estudo da ciência. Entretanto seria um absurdo negarmos que somos conscientes e nos sobra, assim, a segunda alternativa. Mas será que ela é satisfatória? Uma ciência que se proponha a explicar a totalidade da realidade, uma ciência completa, não poderia prescindir da parte subjetiva da mesma pois, conforme Searle (2006, p. 32), “nem toda a realidade é objetiva; parte dela é subjetiva”.

Para iniciarmos então uma análise da proposta de Searle acerca da ontologia da consciência, precisamos apontar para uma confusão identificada por ele sobre a utilização dos termos “subjetivo” e “objetivo”. Tais termos são eventualmente utilizados com sentido epistêmico ou ontológico (SEARLE, 2006, p. 32). Entretanto a distinção entre “subjetivo” e “objetivo” no sentido epistêmico é bem diferente daquela no sentido ontológico. Epistemicamente a diferença entre “subjetivo” e “objetivo” se dá pela tentativa de eliminação de quaisquer aspectos subjetivos em uma afirmação, por parte de quem faz tal afirmação. Explicando melhor: para que o raciocínio não fique circular, epistemicamente, nas palavras de Searle (2006, p. 32), “a distinção discrimina diferentes graus de independência de asserções em relação aos caprichos de valores especiais, preconceitos pessoais, pontos de vista e emoções”. Entretanto, como estamos investigando brevemente a ontologia da consciência, utilizaremos a distinção ontológica entre “subjetivo” e “objetivo” – usada por Searle (2006, p. 32) –, a saber, que a realidade empírica possui duas categorias: objetiva e subjetiva. Então, o sentido do termo “subjetivo” a ser utilizado nesta pesquisa não se refere à subjetividade

epistêmica, mas à ontológica, isto é, o estatuto ontológico do objeto de estudo (a consciência) é que se enquadra categoricamente como subjetivo e não a forma epistêmica de investigação. Searle (2006, p. 139) procura deixar bem claro que “'subjetivo' refere-se a uma categoria ontológica, não a um modo epistêmico”. Assim, a consciência é subjetiva no sentido em que é experimentada por um sujeito (SEARLE, 1998, p. 25).

a consciência não parece ser um fenômeno público igualmente acessível para todos os observadores, mas algo radicalmente privativo do indivíduo cuja consciência ele é<sup>15</sup>, algo incluído a partir de uma perspectiva puramente subjetiva de primeira pessoa. O comportamento das moléculas, por outro lado, é algo público, que pode ser apreendido pelo ponto de vista da terceira pessoa. (MASLIN, 2009, p. 168)

Um estado consciente que é experimentado por um sujeito, só o é pelo sujeito em questão, por mais ninguém: os estados conscientes são considerados, assim, como sendo acessíveis somente em primeira pessoa. Neste sentido Maslin (2009, p. 87) afirma que “embora os cientistas e eu tenhamos igual acesso, em princípio, ao que acontece em nossos cérebros, somente eu tenho acesso aos conteúdos da minha mente”. Podemos dizer mais: os estados conscientes são qualitativos, o que significa que, para qualquer estado consciente, o sujeito se sente como estando em tal estado. Então, mesmo que Owen Flanagan (1991) defenda uma ciência da mente que atente a relatos subjetivos sobre experiências conscientes agregando resultados da psicologia, da ciência cognitiva e da neurociência, os conteúdos qualitativos continuam sendo de acesso exclusivo em primeira pessoa. Portanto, uma dor não pode ser reduzida, por exemplo, ao disparo de fibras C, como defende a corrente da teoria da identidade – na verdade, uma dor é como ela é sentida.

Pelo fato de a consciência possuir uma ontologia de primeira pessoa – por ser ontologicamente subjetiva –, as categorias tradicionais de mental e físico não podem ser

---

<sup>15</sup> Aqui Maslin afirma, de certa forma, que a consciência é o próprio eu. Será necessário – para verificar se tal relação de identidade pode ser aceita – o que se entende por consciência e, também, qual é a concepção de eu.

aplicadas a ela (SEARLE, 1998, p. 26). Ela é qualitativa e subjetiva, mas também é parte natural do mundo físico.

Consciência, em resumo, é uma característica biológica de cérebros de seres humanos e determinados animais. É causada por processos neurobiológicos, e é tanto uma parte da ordem biológica natural quanto quaisquer outras características biológicas, como a fotossíntese, a digestão ou a mitose. (SEARLE, 2006, p. 133)

A consciência para Searle (1998, p. 26) é, assim, parte de uma realidade biológica: é uma característica do cérebro. É um fenômeno natural e biológico que é interno, qualitativo e de primeira pessoa. Ela é causada pelo cérebro em seus microprocessos de nível inferior, mas é uma propriedade desse cérebro em nível “macro” (SEARLE, 1998, p. 26). Ou melhor, os processos neuronais de baixo nível, os microprocessos que acontecem no cérebro, causam os estados conscientes (SEARLE, 2001, p. 272), mas de um modo que ainda não sabemos. A consciência é enfim uma propriedade emergente do cérebro<sup>16</sup>.

Além de qualitativa e subjetiva, a consciência possui outra característica: a de ser unificada (SEARLE, 2001, p. 271). Isto significa que quando estamos tendo um estado consciente, tudo o que experienciamos em tal momento faz parte do campo unificado da consciência. Então, quando assistimos a uma partida de futebol, por exemplo, a sensação de maciez do sofá, o frescor da brisa que vem do ventilador e o sabor salgado da pipoca não são experienciados separadamente. Não há uma consciência da maciez do sofá separada da consciência do sabor salgado da pipoca, por exemplo. O estado consciente é um campo unificado que abarca tudo o que está sendo experienciado: não há como separar, no exemplo anterior, a sensação de maciez do sofá do frescor da brisa que vem do ventilador e do sabor salgado da pipoca, quando são experienciados simultaneamente. Todas estas experiências juntas acabam compondo o campo unificado da consciência. Alguém poderia dizer que, ao provar a pipoca, estava consciente do seu sabor salgado e que tanto a sensação de maciez do

---

<sup>16</sup> “Uma propriedade emergente de um sistema é aquela que é causalmente explicada pelo comportamento dos elementos do sistema [...]”. (SEARLE, 1998, p. 44).

sofá quanto o frescor da brisa que vem do ventilador não estavam presentes à consciência. Searle considera um erro pensar desta forma: para ele a atenção é que – no exemplo – estaria voltada ao sabor salgado da pipoca. Assim, a sensação de maciez do sofá e o frescor do ar que vem do ventilador estariam numa “periferia” da atenção, mas todas as experiências comporiam um único estado consciente (SEARLE, 2010, p. 10-11, 65-66).

Em suma a consciência possui três características fundamentais, ela é qualitativa, subjetiva e unificada. Além disso,

A consciência é uma propriedade mental, e portanto física, do cérebro, no sentido em que a liquidez é uma propriedade de sistemas de moléculas. [...] o fato de uma característica ser mental não implica que não seja física; o fato de uma característica ser física não implica que não seja mental. [...] a consciência *enquanto* consciência, *enquanto* mental, *enquanto* subjetiva, *enquanto* qualitativa, é *física*, e física *porque* mental. (SEARLE, 2006, p. 26)

Tais afirmações demonstram o rompimento com as categorias tradicionais e com todo o vocabulário que elas deixaram como herança. Entretanto tal concepção – apesar de diferente das fisicalistas e funcionalistas – parece, até então, não escapar do comprometimento com o determinismo. Trataremos disso adiante.

#### **4.3 INTENCIONALIDADE: A DIFERENÇA ENTRE A EXPERIÊNCIA DE PERCEBER E A DE AGIR**

O termo *intencionalidade* foi utilizado pelos filósofos medievais da Escolástica e reintroduzido por Franz Brentano no século XIX. Posteriormente Husserl também trabalhou tal conceito a partir das influências recebidas tanto de Brentano quanto de Frege. (PIRES, 2012, p. 287). O conceito de intencionalidade é chave para tratar da questão do livre-arbítrio. Isto porque a ação livre depende de que os estados mentais do agente possam ser sobre o que está além destes próprios estados mentais. Para tornar mais claro o que significa os estados mentais serem sobre algo que esteja além deles próprios, acompanhemos o seguinte exemplo.

Se alguém apenas reagisse à sensação de dor, tirando rapidamente a mão do contato com a panela quente, não diríamos que o livre-arbítrio fizesse parte do acontecido. A dor não é um estado mental intencional porque não é acerca de algo além de si. A dor é sentida: mas não o é sobre algo. Ela pode ser causada por algo externo, mas não é **sobre** o que a causa. Já o amor é sentido acerca de algo que está fora, por assim dizer, do estado mental em si. O amor neste sentido é intencional, a dor não. O sentido de intencionalidade que estamos utilizando então não se refere somente àquelas situações em que dizemos que Fulano teve a intenção de realizar a ação X ou que Fulano não teve a intenção de realizar a ação X. A intenção<sup>17</sup> em fazer ou não fazer algo é apenas um tipo de intencionalidade. Entretanto o sentido de intencionalidade que será utilizado nesta pesquisa é mais amplo: é o de ser uma característica de certos estados mentais, ou seja, destes estados serem sobre o que é externo a eles próprios. Delimitando o conceito de intencionalidade, temos que

'Intencionalidade', da forma como os filósofos usam o termo, refere-se àquele aspecto dos estados mentais segundo o qual estes se dirigem a, ou tratam sobre, ou de estados de coisas no mundo que estão além deles mesmos. [...] Qualquer estado que se dirige a algo distinto dele próprio é um estado intencional<sup>18</sup>. (SEARLE, 2001, p. 34-35)

Sendo esta a definição adotada de intencionalidade (tratada como direcionalidade), Searle (2001, p. 35) lista exemplos de estados intencionais – além das intenções (já citadas) –, a saber, “crenças, temores, esperanças [e] desejos”. Mas por que isto é importante? Por que tratar da intencionalidade se estamos investigando o livre-arbítrio? Porque uma ação livre é baseada em escolhas livres e também porque tais escolhas são necessariamente sobre o que é externo ao próprio estado mental. Ademais, como afirmam<sup>19</sup> Boehner e Gilson (2003, p. 287),

---

<sup>17</sup> A intenção é um tipo de estado intencional e como afirma Donald Davidson (2001, p. 7), “Saber a principal razão por que alguém agiu como agiu é saber a intenção com que a ação foi feita”. Original em inglês. Tradução nossa.

<sup>18</sup> Original em inglês. Tradução nossa.

<sup>19</sup> Tratando de Bernardo de Claraval. A obra é sobre filosofia medieval, mas o exemplo é pertinente e ajuda a elucidar sobre a importância da intencionalidade ao se tratar de livre-arbítrio.

o “‘arbítrio' envolve conhecimento e julgamento” e tanto um quanto o outro devem ser acerca de algo que esteja além do próprio estado mental: ambos implicam intencionalidade<sup>20</sup>.

Outro ponto importante que devemos considerar é o seguinte. Será que se um determinado ser possuir livre-arbítrio, ele deverá necessariamente ser consciente e ainda ser capaz de possuir estados intencionais? Para responder a tal pergunta, primeiramente faz-se necessário verificar se consciência e intencionalidade se diferenciam ou se, pelo menos, estão dissociadas. Neste sentido, Brito (2012, p. 106) expõe, acerca da teoria de Franz Brentano, que não há que se falar em consciência sem intencionalidade, ao menos em seres livres. Já conforme Pires (2012, p. 298), para Husserl “a consciência será sempre intencional”, ou seja, neste aspecto o discípulo seguiu o mestre. Mas será que a existência de consciência exige a existência de intencionalidade? Há uma relação de bi-implicação aqui? Acerca disto, Searle (2002, p. 2-3) nos diz que

Intencionalidade não é a mesma coisa que consciência. Muitos estados conscientes não são Intencionais – por exemplo, um sentimento súbito de exaltação – e muitos estados Intencionais não são conscientes – por exemplo, tenho muitas crenças sobre as quais não estou pensando no momento e nas quais posso nunca ter pensado.

Sendo assim, a consciência não implica intencionalidade nem a intencionalidade implica consciência. Vejamos de forma mais clara tal afirmação. O estatuto do estado intencional é diferente do estatuto daquilo sobre o que ele é. Já nos estados conscientes que não são intencionais, a própria experiência daquilo que se está consciente é o estado consciente em si (como a experiência de dor que é a própria dor). (SEARLE, 2002, p. 3). Temos então que, na teoria searleana, a diferença entre os estados intencionais e a consciência é que os primeiros são necessariamente representacionais com relação a seus objetos. O conteúdo dos estados intencionais possui direção de ajuste e condições de verdade ou de satisfação. (SEARLE, 2013, p. 41). Por outro lado a consciência, como já afirmamos

---

<sup>20</sup> Só se pode conhecer ou julgar o que é externo ao próprio estado mental atual.

anteriormente, é qualitativa, subjetiva e unificada. Entretanto um estado consciente tanto pode ser quanto não ser intencional, ou seja, pela tese searleana pode haver estados conscientes que não são intencionais e estados intencionais não conscientes. Considerando, portanto, que podemos ter estados intencionais que não são conscientes ou estados conscientes que não são intencionais, devemos concordar com a afirmação de Searle de que os termos *consciência* e *intencionalidade* não se referem ao mesmo fenômeno. Então forçosamente temos que discordar da afirmação de Teixeira (2008, p. 52) de que não há como pensar sem ser sobre “*alguma coisa*”. O exemplo do sentimento súbito de exaltação dado por Searle contradiz a afirmação de Teixeira, a não ser que não consideremos um estado consciente como uma forma de pensar. Ou melhor, consideraremos estados conscientes não intencionais como uma das formas do pensar. Assim, apesar de estarmos sempre pensando, nem sempre será sobre alguma coisa – intencionalidade e consciência são diferentes e podem ocorrer separadamente.

Estabelecida a diferença e a dissociação entre intencionalidade e consciência, passemos agora à relação entre a intencionalidade e a racionalidade. O exercício do livre-arbítrio implica o uso da racionalidade que – segundo Searle (2001, p. 22-23) – não deve ser entendida como um conjunto de regras mas, *grosso modo*, como as restrições que constituem os fenômenos da intencionalidade e da linguagem. Seres que não fazem uso da racionalidade não podem romper com as determinações causais impostas pelo instinto ou por outras relações físicas, pois não têm biologicamente acesso a outra forma de motivação para suas ações. Por outro lado a racionalidade implica intencionalidade (e consciência), ou melhor, a racionalidade é interna à intencionalidade – não é uma capacidade cognitiva separada<sup>21</sup>. Isto porque um ser racional não pode ser racional sem que seus estados mentais possam ser acerca do que está além deles próprios. Da mesma forma, para que um ser possa agir baseado em razões, deve ser capaz de ter estados conscientes. Se o uso da racionalidade é essencial para o exercício do livre-arbítrio e se intencionalidade e consciência são necessárias para a

---

<sup>21</sup> Rompimento com o modelo clássico. (SEARLE, 2001, p. 8-9).

racionalidade, então intencionalidade e consciência são essenciais para o exercício do livre-arbítrio. Assim, podemos concluir que a intencionalidade é um ponto fundamental para a questão do livre-arbítrio.

Agora podemos responder à pergunta formulada anteriormente. Temos condições de afirmar que se um determinado ser possui livre-arbítrio, ele deve necessariamente ser consciente e ainda ser capaz de possuir estados intencionais.

Outra tese sustentada por Searle (2007, p. 25) é a de que a consciência é uma vantagem evolutiva, pois é uma “característica biológica superior do cérebro”. Nesta mesma linha devemos concluir que o livre-arbítrio também é evolutivamente uma vantagem. Não parece razoável – conhecendo a teoria darwinista – que tenhamos a sensação de possibilidade de escolha entre alternativas e que tal sensação seja apenas ilusória, isto porque não haveria vantagem evolutiva numa consciência que não pudesse influir no mundo (epifenomenal). Logo, dado que a consciência é um fenômeno natural, forçosamente temos que afirmar o mesmo também da intencionalidade. Faz todo sentido que a consciência seja uma vantagem evolutiva e certamente filósofos de outras correntes da filosofia da mente (fiscalistas e funcionalistas, por exemplo) não discordarão disto. Não é crível que uma característica tão saliente de animais superiores não seja uma vantagem em determinados contextos, ou seja, que não seja uma adaptação vantajosa de tais espécies. Da mesma forma, não faz sentido pensar a questão do livre-arbítrio dissociada do evolucionismo: sem tratá-la como uma característica que traz vantagens à espécie. Então, já que a intencionalidade é indispensável para que haja livre-arbítrio, ela também seria uma vantagem evolutiva.

Searle (2002, p. 50) considera que há duas formas primárias de intencionalidade, a saber, a ação e a percepção. Tais formas de intencionalidade são primárias porque não possuem em suas condições de satisfação a causação intencional<sup>22</sup> – diferentemente das

---

<sup>22</sup> “A causação intencional é qualquer relação causal entre um estado intencional e suas condições de satisfação, na qual ou o estado intencional causa suas condições de satisfação ou suas condições de satisfação causam o próprio estado intencional”. (SEARLE, 2013, p. 45). Original em espanhol. Tradução nossa.

crenças e dos desejos que possuem direção de ajuste<sup>23</sup> mente-mundo e mundo-mente respectivamente. Para o filósofo da Universidade da Califórnia (SEARLE, 2013, p. 40-45) os desejos não são verdadeiros nem falsos, mas possuem condições de satisfação que são o ajuste ou o desajuste do mundo à mente, ou seja, é neste sentido que os desejos possuem direção de ajuste mundo-mente. Já as crenças teriam também condições de satisfação, mas com direção de ajuste da mente ao mundo: em caso de ajuste, a crença seria verdadeira; de desajuste, falsa. Entretanto, como foi dito anteriormente, o autor ainda diferencia duas formas mais básicas de experiência na relação que temos com as coisas do mundo: a ação e a percepção (as duas formas primárias de intencionalidade). (SEARLE, 2013, p. 74-75). No primeiro caso, experimentamos a nós mesmos atuando de modo causal sobre o que nos é externo, ao passo que, no segundo caso, experimentamos a atuação causal do que nos é externo sobre nós. É importante notar que estas formas primárias de intencionalidade não dependem de encontrarmos regularidade no mundo: elas são, por assim dizer, mais “diretas”. De fato a ação e a percepção não requerem que a mente se ajuste ao mundo ou que o mundo se ajuste à mente – o que as torna mais primitivas. Simplesmente sentir-se percebendo é perceber. Quando agimos nos sentimos agindo: não há uma mediação. Estas formas de intencionalidade serão importantes para tratarmos da ação livre.

Searle defende que, na vida, não podemos evitar decidir nem agir e que sentimos que temos alternativas disponíveis quando o fazemos. Realmente não há como discordarmos desta afirmação. Pensemos na situação em que estamos sentados imóveis em uma cadeira sem que nada nos prenda ali. Resolvemos então que – apesar de o Sol estar nos ofuscando os olhos naquela posição – continuaremos sentados ali. Até nesta situação de imobilidade há decisão e o fato de continuarmos imóveis pode ser considerado como uma ação. E é uma ação livre, pois o importante nesta situação é que houve a sensação da disponibilidade de alternativas,

---

<sup>23</sup> Para Searle (2013, p. 40-45) haveria três direções de ajuste, a saber, mente-mundo, mundo-mente e nula. Esta última teria como exemplos emoções tais como a vergonha, o orgulho, o desespero e a alegria.

além de uma intenção-na-ação<sup>24</sup>: componente indispensável do livre-arbítrio. Entretanto nos parece que ainda não teríamos uma explicação satisfatória da ação e da decisão como sendo livres. A tese de Searle é a de que há um *gap*<sup>25</sup> que é característica dessa tomada de decisão e ação. Ele (SEARLE, 2001, p. 62) postula o *gap* e o define da seguinte forma: “O *gap* se dá quando as crenças, desejos e outras razões não são experimentadas como condições causalmente suficientes”. Entretanto para o autor “a volição inclui o *gap* de um jeito que a cognição não inclui” (SEARLE, 2001, p. 70), por isso na ação não sentimos as causas como suficientes para a mesma enquanto na percepção as causas nos parecem sempre suficientes. Ou melhor, na ação voluntária, tal sensação de que as causas da ação são insuficientes para determinar a própria ação é ela própria a sensação de liberdade. (SEARLE, 2001, p. 69). Na situação da percepção, por outro lado, não temos tal sensação de liberdade, pois não temos a sensação de ter alternativas. Aqui há um contraste que se faz evidente: a ação possui a característica de ser livre enquanto a percepção tem um caráter determinista. (SEARLE, 2001, p. 68). Entretanto é importante salientar que há duas formas de ação: a racional e a irracional. Na ação irracional as crenças e os desejos são causalmente suficientes, ao passo que na ação racional não há tal vínculo causal. A ação racional se caracteriza justamente por não ter os desejos e as crenças como causas suficientes. Numa decisão racional para agir, a razão deve ser independente-do-desejo. (SEARLE, 2001, p. 131). Para agir racionalmente o agente deve reconhecer uma razão<sup>26</sup> para tanto. Poderíamos dar como exemplo um compromisso ou uma obrigação dentre outros. Então o agente deve reconhecer um fato externo (SEARLE, 2001, p. 127-128) – aí a importância da intencionalidade – como um motivador e torná-lo um

---

<sup>24</sup> A intenção-na-ação – termo técnico utilizado por Searle – é aquela intenção que se tem quando se realiza uma ação. Em ações não premeditadas temos uma intenção-na-ação mas não temos uma intenção prévia. Wakefield e Dreyfus (1991, p. 262) sintetizam a explicação searleana dizendo que “a ação em si, causada por uma intenção prévia ou não, consiste em um movimento corporal e uma intenção na ação, onde a intenção na ação continua a causar o movimento corporal durante toda a duração do movimento”.

<sup>25</sup> Tal termo foi traduzido por *intervalo* na obra *Liberdade e neurobiologia* (Searle, 2007). Podemos traduzir também por *lacuna*.

<sup>26</sup> Quando temos consciência de que agimos baseados em uma razão, explicar a ação com base naquela razão é perfeitamente adequado. Entretanto isso ainda não explica como pode haver o rompimento com o princípio da causalidade – o que caracteriza o livre-arbítrio.

motivador interno através da representação. Ou seja, o mundo externo ao agente oferece os fatos e estes passam a servir como fundamentos racionais para os seus motivadores internos. É a partir daí que o *gap* se faz presente. O *gap* ocorre entre o desejo motivador e a tomada de decisão. Isto porque a obrigação racional de agir – reconhecida internamente pelo agente – não é causalmente suficiente. Se o fosse, não haveria tomada racional de decisão, pois o caráter da mesma seria determinista.

Alguém poderia argumentar que Searle está confundindo livre-arbítrio e liberdade de escolha, que o determinismo não significa que não podemos fazer escolhas ou que as nossas escolhas não importam, mas que apenas aponta para a existência de razões pelas quais fazemos qualquer escolha. Ou seja, alguém poderia afirmar que Searle está confundindo determinismo com fatalismo e cometendo a falácia do espantalho<sup>27</sup>. Porém não é esta a posição searleana. O que ele defende é que temos uma sensação de liberdade na ação voluntária, ou seja, que temos uma sensação de que as causas da ação são insuficientes para determinar a ação. (SEARLE, 2001, p. 69). E, como afirmamos anteriormente, não é razoável, apesar de ser possível – conforme os nossos conhecimentos acerca da teoria de Darwin –, que tenhamos tal sensação e que ela seja apenas uma ilusão. É o que se evidencia a partir da afirmação de que "[o] epifenomenalismo é uma tese possível, mas é absolutamente inacreditável, e se a levássemos a sério, mudaria a nossa visão de mundo [...]". (SEARLE, 2001, p. 286). Assim, a posição searleana procura dar conta de como a ação voluntária pode romper com o princípio da causalidade.

Enfim, sintetizando *grosso modo* a importância da questão da intencionalidade na questão do livre-arbítrio, podemos dizer o que segue. A ação livre é baseada em escolhas livres. Estas se dão sobre o que é externo aos próprios estados mentais, ou seja, é constitutivo de tal escolha a intencionalidade intrínseca: a ação livre depende de intencionalidade. Já o

---

<sup>27</sup> A falácia do espantalho é uma forma de argumentar atacando uma posição que o interlocutor não defende. Ela pode ser proposital ou não.

arbítrio envolve tanto conhecimento quanto julgamento, os quais também só podem ser sobre aquilo que é externo aos próprios estados mentais. Mas o julgamento depende da racionalidade a qual é interna à intencionalidade, ou melhor, a racionalidade oferece restrições aos fenômenos intencionais. (SEARLE, 2001, p. 108). Por isso a ação livre pressupõe a racionalidade. Mas a ação racional parte de que crenças e desejos não sejam causalmente suficientes para que ela ocorra. Se as crenças e os desejos – que são formas não primitivas de intencionalidade – estiverem envolvidas causalmente como motivadores da ação, poderemos afirmar que tal ação é irracional. Isto porque a racionalidade já pressupõe a livre escolha (SEARLE, 2001, p. 142). Ora, se a racionalidade fosse inerte (epifenomênica), a deliberação não teria lugar na ação humana e, não havendo deliberação, as crenças e os desejos seriam causalmente suficientes para a ação. Quando as crenças ou os desejos são suficientes para a ação, esta não é livre. Assim, como podemos encontrar nas palavras de Aristóteles (2002, p.83; EN 1110a), “as ações são involuntárias quando realizadas sob compulsão ou através da ignorância”. Entretanto, por enquanto, teríamos uma relação de causa e efeito: as causas da ação e o efeito, sendo este último em forma de ação. Ou melhor, na relação de causa e efeito, as causas da ação fariam o “papel” de causa, enquanto que a ação seria o efeito. Mas se fosse puramente de causa e efeito, a relação seria de caráter determinista. Como a nossa experiência ao agir é a de que temos alternativas, deve haver algum componente a mais nesta relação. Assim, entre as causas e o efeito teríamos um *gap* ao qual Searle (2001, p. 13) identifica como sendo o livre-arbítrio.

Mas o que torna a ação livre? O *gap*? E o que isto tem a ver com a intencionalidade? Searle (2001, p. 65) defende que quando efetuamos um movimento corporal voluntariamente, a intenção-na-ação é causalmente suficiente para que tal movimento se realize. Para aquele filósofo não é o movimento físico que não tem causas suficientes, mas sim a intenção-na-ação. A intenção-na-ação, portanto, não possui causas suficientes – nem físicas nem

psicológicas. Temos aí então a manifestação do *gap* na liberdade humana. Mas a intenção-nação é uma forma de intencionalidade. Portanto, a ação do agente livre, enquanto tal, depende de que seus estados mentais possam sobre o que é externo a eles próprios. Por conseguinte, a ação livre pressupõe intencionalidade intrínseca.

#### 4.4 NATURALISMO BIOLÓGICO *VERSUS* FISCALISMO

Acabamos de ver como a questão da intencionalidade intrínseca tornou-se o núcleo para uma análise sobre a possibilidade ou não do livre-arbítrio. Julgamos que a conclusão da seção anterior é plausível porque não há como agir de forma livre se os desejos ou as crenças forem causalmente suficientes. Ela nos será útil para a análise comparativa do naturalismo biológico de John Searle com o fiscalismo em relação à questão do livre-arbítrio.

Começaremos por explicitar a diferença entre o físico e o material para tornar claro com qual corrente o naturalismo biológico searleano está sendo comparado criticamente. Posteriormente partiremos de que a intencionalidade intrínseca daquele que realiza a ação é necessária para que haja livre-arbítrio e verificaremos se as correntes do fiscalismo são tão ou mais compatíveis com tal noção do que o naturalismo biológico de Searle. Como resultado desta análise, pretendemos nesta seção identificar qual é o melhor caminho a se seguir numa investigação acerca do livre-arbítrio: a posição searleana ou a fiscalista.

Para distinguir o físico do material, vejamos o que segue. Fiscalistas entendem que a realidade pode ser explicada unicamente através da linguagem da Física. Assim, não haveria nada privado. Mesmo os estados mentais subjetivos – os *qualia* – não seriam de acesso exclusivo em primeira pessoa. Na corrente fiscalista da filosofia da mente, tudo que é físico tem propriedades que são ou físicas ou estreitamente relacionadas com o que é físico. Já o materialismo se enquadra como uma forma de fiscalismo. A diferença para as outras

modalidades de fisicalismo é que para os materialistas somente há a matéria como substância, ou seja, somente da matéria se pode dizer que existe. Como a Física já nos deixou claro que no Universo há tanto matéria quanto energia, as quais são diretamente proporcionais, conforme a célebre fórmula de Einstein<sup>28</sup>, consideramos mais adequado tratar do fisicalismo. Assim, podemos dizer, de certa forma, que o que é material está contido no físico. De fato, campos magnéticos são fenômenos físicos ao mesmo tempo em que não são constituídos de matéria.

Dadas tais considerações, devemos desconsiderar posições exclusivamente materialistas. Então estamos autorizados a fazer a análise comparativa do naturalismo biológico com modalidades de fisicalismo que concebem a existência não somente de matéria<sup>29</sup>.

Consideraremos a teoria da identidade mente-cérebro como a principal representante do fisicalismo. Mas o que ela defende? A teoria da identidade é contra o dualismo de substâncias e defende que a mente e o cérebro são a mesma coisa, isto é, que quando falamos da mente estamos falando do cérebro e vice-versa.

Passemos então à análise comparativa da posição searleana com as correntes do fisicalismo. Como já foi dito, a intencionalidade intrínseca é o ponto central para esta discussão e será o nosso fio condutor para a análise acerca da possibilidade do livre-arbítrio nas referidas correntes da filosofia da mente. Uma estratégia que pretendemos utilizar é a de primeiramente verificar os pontos comuns e, posteriormente, as divergências – se houver – entre o fisicalismo e o naturalismo biológico.

Tanto as correntes do fisicalismo quanto John Searle pretendem que as suas posições não carreguem nenhum traço de dualismo, sobretudo o de substâncias. Este é um ponto em comum entre as duas posições: ambas rechaçam o dualismo ontológico. Entretanto o

---

<sup>28</sup>  $E=mc^2$ .

<sup>29</sup> Não estamos falando de um fisicalismo dualista no que se refere a substâncias, mas de uma posição que concebe a existência de fenômenos e propriedades.

naturalismo biológico searleano é acusado por Prata (2013, p. 269) de alinhar-se a um dualismo de propriedades. O argumento é que seria possível conceber que um determinado processo cerebral pudesse causar um *qualium* diferente daquele que no mundo atual ele causa. Assim, a relação entre processos cerebrais e estados qualitativos seria contingente. Discordamos de tal posição porque os mundos possíveis seriam mundos “como se”. Em algum momento de uma análise mais aprofundada de tal mundo, chegaríamos a uma impossibilidade lógica ou nomológica. Ou melhor, se nos dedicássemos a “dissecar” um mundo “possível” em que tudo é exatamente igual ao nosso mundo com a exceção de que não há consciência<sup>30</sup>, não há *qualia*, chegaríamos – em algum momento – a alguma impossibilidade. Nosso ponto de vista é o mesmo de Searle (1998, p. 164) quando ele diz que “um caso como este é impossível pois sabemos que a estrutura e a função do cérebro são causalmente suficientes para produzir a consciência”. Kim (1998, p. 101) também afirma ser “inconcebível que exista um mundo possível que seja uma duplicata física exata deste mundo, mas que a intencionalidade esteja totalmente ausente”. Se estados intencionais podem influir no mundo e se o contrário também ocorre, não há como haver um mundo exatamente como é sem que um aspecto dele permaneça. Os estados intencionais fazem parte do mundo – assim como a consciência – e negar isso nos dá apenas duas alternativas, a saber, o eliminativismo e o epifenomenalismo. Ou melhor, se afirmarmos que a subjetividade não faz parte da realidade, ou nós a eliminamos ou a tratamos como se estivesse fora do mundo (o que é uma posição de certa forma dualista). Mas, desta forma, dentro de uma concepção de um mundo causalmente fechado, fisicamente falando, a subjetividade seria epifenomenal.

É importante observar que a possibilidade de realização múltipla, em tese, é possível: dois cérebros em configurações diferentes poderiam ter *qualia* idênticos. Entretanto, dois cérebros em configurações exatamente iguais até a última partícula, jamais poderiam apresentar aspectos qualitativos e subjetivos diferentes. Searle (2006, p. 180) reforça esta

---

<sup>30</sup> Este é o famoso argumento dos zumbis. (KIM, 2008, p. 27).

diferença quando afirma que “identidade de neurofisiologia garante identidade de mentalidade; contudo, identidade de mentalidade não garante identidade de neurofisiologia”.

Acabamos de afirmar que negar que os estados intencionais fazem parte do mundo nos coloca entre o eliminativismo e o epifenomenalismo. Mas além disso, baseando-se num dilema apresentado por Kim (1998) aos fisicalistas, Bensusan (2002) defende que o fisicalismo é insustentável. O dilema consiste em que um fisicalista ou deve aceitar o epifenomenalismo ou deve aceitar que propriedades mentais são propriedades físicas. Assim, ou o mental não tem qualquer influência sobre o físico, ou o mental é físico. O próprio Kim procura fugir da primeira opção, até porque não vê muita diferença entre o epifenomenalismo e o eliminativismo (KIM, 1998, p. 119). De fato, algo que não exerce influência no mundo, causalmente falando, existe?

Havíamos tratado anteriormente do conceito de superveniência<sup>31</sup> sem nos aprofundarmos muito. Tal conceito é assumido por Jaegwon Kim parcialmente. Para ele há superveniência entre propriedades (físicas e mentais). (KIM, 1998, p. 9). Mas se o que é subveniente – que serve de base – não pode ser alterado pelo que é superveniente, o fisicalismo que aceita a superveniência acaba por desembocar outra vez no epifenomenalismo – o que tem implicações sérias na questão do livre-arbítrio. Novamente: se o mental é epifenomenal, não tem poder causal algum – o que vai contra o que sabemos de biologia sobretudo de evolução das espécies. Ora, se o conteúdo semântico dos estados mentais não possui poder causal e se a intenção-na-ação é fator imprescindível para a ação livre, há de se concluir que um fisicalismo assim concebido de fato ou não consegue dar conta do livre-

---

<sup>31</sup> David Chalmers (1996) aborda a superveniência de forma mais aprofundada. Ele nos diz o seguinte sobre a definição de superveniência: “Em geral, a superveniência é uma relação entre dois conjuntos de propriedades: propriedades-B – intuitivamente, as propriedades de alto nível – e propriedades-A, que são as propriedades de baixo nível mais básicas. [...] As propriedades-A relevantes são geralmente as propriedades físicas: mais precisamente, as propriedades fundamentais que são invocadas por uma teoria completa da Física. [...] Os fatos-A e os fatos-B sobre o mundo são os fatos relativos à instanciação e distribuição de propriedades-A e propriedades-B. [...] O modelo para a definição de superveniência é o seguinte: propriedades-B sobrevêm em propriedades-A se não há duas situações possíveis idênticas no que diz respeito às suas propriedades-A, enquanto diferem em suas propriedades-B.”. (CHALMERS, 1996, p. 33).

arbítrio ou o nega porque atribui toda a suficiência causal aos fatos físicos brutos. Neste sentido Searle (2006, p. 182) argumenta contra a noção de superveniência afirmando que ela é inútil para a Filosofia. Concordamos com Searle primeiramente pelo motivo já exposto, ou seja, que tal noção vai contra o que há de mais consolidado no darwinismo. Ora, se o mental fosse superveniente, qual seria a sua vantagem evolutiva? Ademais Searle também defende que a causalidade micro-macro já nos dá o caminho para explicar como tal relação acontece.

Minha conclusão é que, quando você reconhece a existência de formas de causalidade de baixo para cima, micro para macro, a noção de superveniência deixa de ter qualquer função na filosofia. As características formais da relação já estão presentes na suficiência causal das formas de causalidade micro-macro. (SEARLE, 2006, p. 182).

Alguns fisicalistas como Jaegwon Kim utilizam a noção de superveniência para tentar fugir da admissão da existência de entidades imateriais ou então para não assumir formas de ser que não sejam naturais. Searle não segue este caminho. Como ele não aceita as categorias tradicionais – o que é um ponto de divergência com as posturas fisicalistas –, não necessita da noção de superveniência. Prata (2012) concorda que há divergências entre o naturalismo biológico de Searle e o fisicalismo. Sobretudo quando o primeiro defende que “nem toda a realidade é objetiva”, ontologicamente falando, mas que “parte dela é subjetiva”. (SEARLE, 2006, p. 32). Isto porque, segundo ele, “[a] adesão de Searle a esse dualismo entre subjetividade e objetividade afasta o seu naturalismo biológico do fisicalismo [...]”. (PRATA, 2012, p. 136).

Até agora temos que posições que têm de se ancorar na superveniência ou no epifenomenalismo para explicar o mental não oferecem razões satisfatórias que nos façam dizer que elas sejam o caminho para uma investigação acerca do livre-arbítrio. Tratamos anteriormente também da incompatibilidade de posições eliminativistas com o livre-arbítrio pois, ao reduzir os conteúdos e estados mentais ao físico (dentro de um paradigma de terceira

pessoa), como o faz Churchland, segundo Cescon (2013, p. 94), ou elimina o mental ou o torna sujeito às leis deterministas da Física.

Uma outra posição physicalista em filosofia da mente que podemos comparar criticamente com a de Searle é a de Donald Davidson por sua importância no debate filosófico contemporâneo. Seu pensamento se aproxima em parte ao de Searle, principalmente ao rechaçar o dualismo de substâncias. Davidson (2001) defende sobretudo duas teses em sua filosofia da mente, sendo que a primeira é a de que eventos mentais e eventos físicos são idênticos entre si. A outra tese é a de que eventos mentais não são regidos por leis estritas, ou seja, que o mental é anômalo. Destas duas teses principais surgiu o nome de sua posição no âmbito da filosofia da mente, ou seja, o monismo anômalo. Segundo Sparano (2003, p. 168), é justamente por não existir – pela visão de Davidson – leis precisas que façam a ligação do mental com o físico que tal posição é contra o reducionismo psicológico. Klaudat (1999, p. 48) reforça que “Davidson *não deseja* ser um reducionista” dos eventos mentais a processos cerebrais. Davidson pode ser classificado como um physicalista por não aceitar o reducionismo ontológico; apenas o linguístico. Para ele (DAVIDSON, 2001) o mental e o físico são formas (irreduzíveis uma à outra) de explicar e entender as relações causais entre eventos. Por outro lado, Davidson considera que todos os eventos são físicos (SPARANO, 2003, p. 169), pois eventos mentais são idênticos a eventos físicos (KLAUDAT, 1999, p. 53). Então, sendo todos os eventos físicos, o que diferenciaria os mentais dos demais é possuir, adicionalmente às características físicas que todos têm, características mentais. Sparano (2003, p. 208) procura esclarecer este ponto do pensamento davidsoniano da seguinte forma.

Os eventos têm vários movimentos, mas a ação é única, ligada por um encadeamento causal. Davidson quer mostrar como a teoria, que tem na sua base a noção de causa, “o cimento do universo”, é compatível com a ideia de que a liberdade de agir é um poder causal sem ser determinista ou restrito à causalidade original ao agente. O sujeito passa a ser um produto de suas ações, mais do que a fonte última da intencionalidade que determina seu comportamento.

Conforme Sparano (2003, p. 179), Davidson (2001) defende que a ação livre é baseada em crenças, intencionalidade e racionalidade. Entretanto, para ele o sujeito da ação não é aquele que causa, mas o que é um produto dela. Tal posição parece estar mais alinhada com a de Hume (*Tratado*, Livro I, Parte IV, Seção VI) – de que não há um “eu” – do que com a de Searle (2013, p. 93) que defende que “a ação pressupõe um eu”.

Segundo Sparano (2003, p. 171), a racionalidade – na teoria de Davidson – tem como função “dotar de sentido as atitudes e decisões intencionais do agente”. Entretanto, confrontando tal posição com o naturalismo biológico searleano, defendemos que a racionalidade é apenas uma parte da intencionalidade e que mesmo que determinadas atitudes ou decisões não sejam racionais elas não são desprovidas de sentido. Ora, a atitude de Fulano ao dar um tapa em seu irmão, movido pela raiva, pelo fato de ele ter pego o seu caminhãozinho de brinquedo, pode ser uma ação irracional – por não ser uma ação independente-do-desejo – mas não deixa de possuir sentido.

Em suma, podemos afirmar que há uma grande diferença entre as posições filosóficas searleana e davidsoniana acerca do livre-arbítrio. Podemos dizer que, para Davidson (2001, p. 3-19), as raízes causais da ação livre estão no início das crenças e dos desejos, enquanto que para Searle (2013, p. 169) a ação livre não pode ter condições causalmente suficientes – se for este o entendimento de causa. Para Searle, a ação livre é independente-do-desejo.

Ora, parece coerente o raciocínio de Searle. Se a ação é causada por crenças ou desejos, podemos dizer que estas são causalmente suficientes para aquela. Porém, se assim for, não há escolha: aquele que realiza a ação de fato não tem outra alternativa, pois também não há margem para deliberação. Mas convenhamos, se é assim, a racionalidade é inerte (ou epifenomenal) e a ação, por sua vez, fica fora do alcance da valoração racional. (SEARLE, 2013, p. 153-154).

Davidson, juntamente a muitíssimos outros filósofos, pensa que no caso das ações motivadas racionalmente existe algum tipo de conexão causalmente necessária entre os antecedentes psicológicos da ação e sua realização intencional, ou ao menos a

pretensão intencional de realizar a ação, de tal forma que a ação se segue de seus antecedentes por algum tipo de necessidade causal. Porém isto é um erro. Com ele a existência do gap<sup>32</sup> é negada. (SEARLE, 2001, p. 229).

Ações que possuem causas psicológicas suficientes (SEARLE, 2001, p. 230) são inevitáveis do ponto de vista daquele que age. Portanto, se concordarmos que as ações livres não podem ter causas suficientes, deveremos concluir que o naturalismo biológico de Searle está mais condizente com o que entendemos por livre-arbítrio.

#### 4.5 NATURALISMO BIOLÓGICO *VERSUS* FUNCIONALISMO

Um ponto central que defendemos até aqui é que a intencionalidade intrínseca deve ser o fio condutor para a pesquisa acerca do livre-arbítrio. Já havíamos visto também que o funcionalismo é uma espécie de fisicalismo e que, por isso, padece das mesmas dificuldades quanto à questão do livre-arbítrio. Entretanto, por ser uma corrente de pensamento muito difundida, o funcionalismo merece ser analisado. Até porque talvez as suas especificidades – que o diferenciam das outras correntes fisicalistas – possam manter tal posição como uma candidata respeitável no que se refere à questão do livre-arbítrio. Portanto, assim como foi dito anteriormente, faremos a comparação crítica do naturalismo biológico com o funcionalismo tendo como enfoque a questão da intencionalidade intrínseca articulada com a do livre-arbítrio.

Retomemos a caracterização do funcionalismo. Churchland (2004, p. 68) explica que o funcionalismo pode ser confundido com o behaviorismo. E até este ponto da nossa discussão, esta confusão pode ter aparecido de forma sutil. Churchland (2004) nos diz que a diferença entre o funcionalismo e o behaviorismo é que para este os estados mentais podem ser definidos exclusivamente em termos de entrada de dados e saída (em forma de comportamento). Já para os funcionalistas os estados mentais só podem ser definidos

---

<sup>32</sup> Já vimos anteriormente que para Searle é a intenção-na-ação que é causalmente suficiente na ação livre. Vimos também que é a intenção-na-ação que não possui causas suficientes. O livre-arbítrio, para Searle, está no *gap* entre as causas e o efeito.

relacionado-os causalmente com outros estados mentais. Segundo Maslin (2009), para os funcionalistas – que caracterizam a mente como uma função – termos mentais ou psicológicos são, em princípio, elimináveis de uma certa maneira. Se dissermos que nosso vocabulário mental consiste de nomes para estados e relações mentais, a alegação poderá ser de que esses nomes podem ser tratados como sinônimos de descrições definidas, cada uma das quais pode ser formulada, em princípio, independentemente do uso de qualquer vocabulário mental. Ou seja, de certa forma o funcionalismo converge para o fisicalismo e elimina a subjetividade. Mas qual seria o ponto de divergência entre o funcionalismo e o fisicalismo? Segundo Churchland (2004, p. 68-69), a diferença reside no fato de os fisicalistas identificarem o mental com o físico, enquanto os funcionalistas concebem que o mesmo estado mental pode ocorrer em uma base física totalmente diferente, desde que execute a mesma função. Não parece razoável tal distinção, pois o fisicalismo – se não for o caso de implicar um eliminativismo – concebe que o mesmo estado mental pode ocorrer em configurações diferentes do cérebro ou de outras formas de organização do mundo físico. É por isso que Churchland (2004, p. 69) relaciona o funcionalismo com o fisicalismo espécime-espécime, diferencia-o do fisicalismo tipo-tipo. Entretanto o funcionalismo tem dificuldade em explicar o mental principalmente porque considera que a organização funcional é que faz com que ele seja o que é. Ou seja, qualquer base física poderia ter mente desde que tivesse uma organização funcional que permitisse isso. Searle (1998, p. 160) então indica que os funcionalistas se encontram num dilema: “desistir do funcionalismo e aceitar a irreduzibilidade da consciência ou ficar com o funcionalismo e negar a irreduzibilidade da consciência”. Mas o que coloca os funcionalistas contra a parede, segundo Searle? O que os leva a tal dilema? Searle (1998, p. 160), para colocar o funcionalismo em xeque, utiliza o exemplo das dores, ou seja, de que elas (conforme a concepção de senso comum) são (i) “internas, qualitativas e subjetivas”, bem como são (ii) “causadas por processos

neurobiológicos específicos no cérebro e no resto do sistema nervoso”. O problema, segundo Searle (1998, p. 160), é que, ao afirmar que dores são “estados físicos que são partes de uma organização funcional nos cérebros ou em algum outro sistema”, os funcionalistas são forçados a negar (i). Da mesma forma, quando afirmam que “estados físicos, funcionalmente organizados, não causam dores”, mas que “eles apenas são dores”, necessariamente têm de rejeitar (ii). Mas qual é o problema disso? O problema é que para salvar a sua teoria, os funcionalistas têm de negar a consciência, os *qualia*, a subjetividade e – por conseguinte – a intencionalidade intrínseca.

Um autor ao qual Searle (1998) acusa de negar a consciência é Daniel Dennett. Dennett (1997) defende que a nossa intencionalidade – que Searle chama de intrínseca – é derivada de outros sistemas mais primitivos que também possuem intencionalidade derivada de outros sistemas e assim sucessivamente. Isto fica claro no seguinte trecho.

[...] a intencionalidade que nos permite falar e escrever e nos maravilhar com todas as maneiras de maravilhar-se é inegavelmente um produto complexo e tardio de um processo evolutivo que possui os tipos mais grosseiros de intencionalidade – menosprezados por Searle e outros como 'simples intencionalidade do tipo como se' – como ancestrais e componentes contemporâneos. Somos descendentes de robôs, e compostos por robôs, e toda a intencionalidade que desfrutamos é derivada da intencionalidade mais fundamental desses bilhões de sistemas intencionais grosseiros.(DENNETT, 1997, p. 55).

Ora, Dennett parece ignorar que a intencionalidade intrínseca (ou original) pressupõe subjetividade. Quando diz que a intencionalidade dos nossos estados mentais é simplesmente intencionalidade derivada, aparentemente exclui a subjetividade do mental, a consciência. Dennett parece afirmar que a mente humana constitui-se de um aglomerado de sistemas brutos de intencionalidade derivada. Até aqui a nossa posição é a mesma de Searle, ou seja, a de que mesmo os estados intencionais não conscientes – como as crenças das quais não estamos conscientes neste momento – são estados de um *eu* que possui estados conscientes. Por isso nos parece um tanto estranha a afirmação de Dennett acima: a nossa crença de que lá fora está chovendo não seria (conforme ele) um exemplo de intencionalidade intrínseca, mas

de uma intencionalidade derivada de outras intencionalidades derivadas. Se seguirmos por tal caminho teremos duas consequências, a saber, um regresso ao infinito e a eliminação da consciência. A primeira se segue de que, se cada forma de intencionalidade for derivada de outra, teremos uma série infinita de intencionalidades derivadas. A segunda vem da negação da intencionalidade intrínseca: alguns estados conscientes são intencionais; e são intrinsecamente intencionais, como uma crença consciente. Negar a intencionalidade intrínseca é também negar a consciência ou se comprometer com o epifenomenalismo. Como já havíamos tratado anteriormente, o epifenomenalismo pode até ser possível (SEARLE, 2013, p. 306), mas não é compatível com o conhecimento científico sobre biologia e evolução das espécies.

Dennett supõe que a sua posição – conforme a citação supra – seja a mais científica, por assim dizer, no âmbito da filosofia da mente. Para ele, memes<sup>33</sup> se propagariam através da intencionalidade derivada. De fato Dennett (1991, p. 210) defende que “a consciência humana é, ela mesma, uma imensa coleção de memes”. Mas tal filósofo cai em uma armadilha criada por ele próprio, ou seja, acaba atribuindo intencionalidade a unidades de “imitação” (os memes). Entretanto, como Searle (1998, p. 124) expõe, a teoria darwinista colocou por terra qualquer explicação que atribua teleologia ou intencionalidade tanto à origem quanto ao desenvolvimento das espécies, dentre elas a humana. Quando Dennett adere às “unidades de imitação”, parece esquecer que a imitação pressupõe tanto a consciência quanto a intencionalidade intrínseca. (SEARLE, 1998, p. 124-125). Podemos sempre perguntar “quem imita?”. Se não pudermos atribuir a imitação a um “eu”, atribuiremos a seres brutos que não possuem intencionalidade intrínseca. Isto é contraditório. Desta forma somos impelidos a concordar com Searle acerca da contradição das afirmações de Dennett de que somente há intencionalidade derivada.

---

<sup>33</sup> Segundo Blackburn (1997, p. 243) é um termo criado por Richard Dawkins em 1976 em sua obra chamada “The Selfish Gene”. É uma unidade de evolução cultural passível de replicação, imitação ou autopropagação.

A pretensão de maior cientificidade por parte de Dennett é proveniente também da sua concepção de que a ciência deve ter o ponto de vista de terceira pessoa e de que somente o que pode ser verificado por ela existe. Ou seja, só existe o que está ao alcance ou pode ser verificado pelo ponto de vista de terceira pessoa, portanto a consciência – para ele – não existe. (SEARLE, 1998, p. 131-132). Entretanto, se for o caso, se o método não alcança o objeto de estudo (a consciência, por exemplo), isto não implica a sua inexistência. Não é o objeto de estudo que deve se adaptar ao método, mas o contrário. A consciência tem um estatuto ontológico de primeira pessoa e se o método científico não consegue explicá-la, isso não significa que inexista ou que seja apenas análoga a um programa de computador implementando funções (funcionalismo).

O funcionalismo de Dennett fica evidente em sua afirmação de que “O que torna alguma coisa uma mente (ou uma crença, uma dor ou um temor) não é aquilo de que é feita, mas o que ela pode fazer. Consideramos este princípio incontroverso em outras áreas, especialmente na nossa avaliação de artefatos.” (DENNETT, 1997, p. 65). É um funcionalismo com aspectos behavioristas: algo tem mente simplesmente pelas respostas que dá aos *inputs*, pelo comportamento que apresenta. Desconsidera totalmente a subjetividade e, juntamente, a consciência e a intencionalidade intrínseca. Para ele é desta forma que deve ser uma posição científica e isso nos assusta por não vermos nela a possibilidade do livre-arbítrio.

Nos assusta (sic) a ideia de que a ciência descubra, esteja por descobrir ou ameace com descobrir que nunca seremos o que queremos ser. A ameaça não provém do determinismo – se fosse assim, poderíamos suspirar aliviados pois os físicos estão, aparentemente, de acordo sobre a natureza indeterminista do mundo – senão da própria ciência e da sua visão das coisas: o 'naturalismo'. (DENNETT, 1992, p. 193).

Ou seja, Dennett se considera um naturalista pois, segundo ele, tem a visão científica do funcionamento do mundo e – para o senso comum – a ciência é “inimiga” do livre-arbítrio. Para ele, a sua própria postura é mais científica que a de Searle, por exemplo, porque considera o método de terceira pessoa utilizado pela ciência como o mais adequado ao estudo

da consciência. Segundo ele, o senso comum está errado em ver determinismo na ciência porque ela própria possui indeterminações ou aleatoriedades. Mas tais aleatoriedades relacionadas – por exemplo – à física quântica não podem, segundo Searle (2007, p. 67), ser a fonte do livre-arbítrio, pois liberdade não é o mesmo que aleatoriedade. De fato o próprio Dennett (2006, p. 388) concorda que a aleatoriedade é uma ilusão: o que nos dá a impressão de que certos processos são aleatórios é, segundo ele, a não identificação de um padrão – o que corrobora a posição searleana quanto à não identificação da liberdade com a aleatoriedade. Ademais a hipotética aleatoriedade da decisão consciente e portanto intencional tiraria justamente toda a importância evolutiva da consciência e da capacidade de intencionalidade intrínseca.

Da mesma forma que a aleatoriedade não pode ser a candidata a identificar-se com o livre-arbítrio, a sintaxe não pode ser traduzida por intencionalidade. Talvez possamos dizer que os programas de computador possuem intencionalidade derivada, utilizando os termos de Dennett, mas não há como afirmar que eles possuem intencionalidade intrínseca. Isto porque não sabemos como o nosso cérebro produz a consciência. Não sabemos portanto como surge a nossa capacidade de intencionalidade intrínseca. Mas os estados mentais – diferentemente dos programas de computador – caracterizam-se pelo seu conteúdo. Ou seja, há uma diferença fundamental aqui. Programas de computador – os quais executam funções – são caracterizados por serem formais, não pelo seu conteúdo. Já os estados mentais são definidos pelo seu conteúdo. Mas qual é a diferença afinal? Ora, como Searle (2010, p. 17) afirma, “a sintaxe em si não é suficiente para o conteúdo semântico”. Se a sintaxe não é suficiente para o conteúdo semântico, não é suficiente para a intencionalidade intrínseca.

Tendo em vista a posição dennettiana de somente atribuir intencionalidade derivada aos nossos estados mentais e a postura funcionalista em geral de atribuir o estatuto ontológico de função a tais estados, ficaríamos obrigados a não reconhecer um “eu” como aquele que

realiza a ação ou recairíamos num dualismo. Primeiramente, por que recairíamos em dualismo? Ora, se os funcionalistas defendem que mentes nada mais são que funções sendo implementadas e que uma função não é nem mental nem física (assumindo aqui tais categorias), poderíamos concluir que o funcionalismo defende que as funções pertencem a uma espécie de “terceiro reino”. Isto porque as categorias searleanas de ontologicamente subjetivo ou objetivo não são assumidas pelos funcionalistas. Assim, teria o funcionalismo uma posição platônica como a de Frege<sup>34</sup> acerca das funções? Qual seria o estatuto ontológico das funções? Qual seria o estatuto ontológico dos projetos? Seriam entidades platônicas ou históricas? Em tais questões fica evidente o dualismo ao qual o funcionalismo tende a aderir. A intencionalidade intrínseca neste caso até teria lugar se não assumirmos que os nossos estados mentais possuem somente intencionalidade derivada. Mas a intencionalidade intrínseca seria originada em qual reino? No “terceiro reino”? Não parece ser uma posição compatível com o pensamento científico.

Mesmo que o funcionalista não aceite o dualismo, acaba desembocando em uma posição de não reconhecimento de um “eu” que realiza a ação. Isto porque a posição dennettiana – referida anteriormente – de somente atribuir intencionalidade derivada aos estados mentais implica não reconhecer um ponto de vista de primeira pessoa. Desta forma não reconhece a existência de um “eu”. Entretanto, como afirma Searle (2013, p. 93), um “eu” é pressuposto em toda ação. Para ele uma ação racional é realizada por um “eu” que cria razões independentes-do-desejo para si mesmo. (SEARLE, 2013, p. 217).

No caso da realidade consciente, nada preenche o *gap*. Uma pessoa toma uma decisão e então simplesmente age. Estes fatos nos são inteligíveis somente se postulamos um agente racional consciente, capaz de refletir sobre suas próprias razões e então agir com base em tais razões. [...] [T]al postulação equivale à

---

<sup>34</sup> Gottlob Frege (1998) diferencia o pensamento do pensar. Define que o pensamento não é nem sensível nem pertencente ao “mundo interior” (não é uma representação). Para ele o ser pensante é portador do pensar – é portador de representações – e não de pensamentos. “Se chamamos aquilo que emerge em nossa consciência de representação, então vivenciamos propriamente apenas representações, mas não as suas causas”. (FREGE, 1998, p. 198).

postulação de um eu. Somente podemos entender a ação consciente, livre e racional se postulamos um eu consciente. Mas tal postulação só tem sentido em relação ao fato de um campo consciente unificado de subjetividade. Você não pode dar conta do eu racional apenas nos termos de um feixe humeano de percepções desconectadas. (SEARLE, 2001, p. 288).

A ação consciente, livre e racional é incompatível com a postura funcionalista de restrição à sintaxe. Tal postura está mais próxima das relações de causa e efeito do fisicalismo. Ora, ou o funcionalista identifica o “eu” com função ou nega a existência desse “eu”. Identificar o “eu” com função tem a consequência anteriormente descrita – que elimina a possibilidade de livre-arbítrio. Por outro lado, negar a existência de um “eu” impossibilita aceitar que exista um ponto de vista de primeira pessoa, ou seja, nos força a negar que uma parte da realidade seja subjetiva. Mas porquê? A intencionalidade intrínseca implica um ponto de vista de primeira pessoa, um campo unificado de consciência, um “eu”. Mas não podemos postular um “eu” desprovido do que é fundamental a ele. Por isso não podemos afirmar que um determinado “eu” particular possa ser “implementado” em outra base física. A base física particular original é parte desse “eu” particular. Notemos aqui a distinção entre o “eu” e uma função. A mesma função pode ser implementada em bases físicas diferentes, já o “eu” deixa de existir sem o que é físico (a não ser que assumamos o dualismo de substâncias). Assim, mesmo que o funcionalista assuma a existência do “eu”, se ele considera o cérebro como algo separado desse “eu”, assume, conforme Cinara Nahra (2014, p. 185), um dualismo de substâncias ingênuo. Para ela “as concepções sobre o livre arbítrio mais razoáveis não pressupõem que o livre arbítrio exista independentemente do nosso cérebro”. O cérebro é constitutivo do “eu”. E o “eu” é um postulado necessário para a intencionalidade intrínseca.

Ora, já havíamos concluído que, para que a ação seja livre, o agente livre deve, ao realizá-la, valer-se de intencionalidade intrínseca. Funcionalistas dão ênfase ao aspecto formal e deixam de lado a intencionalidade intrínseca. Sendo assim, o naturalismo biológico

searleano apresenta-se como um caminho mais viável que as correntes funcionalistas nas pesquisas acerca do livre-arbítrio.

## 5 IMPLICAÇÕES ÉTICAS

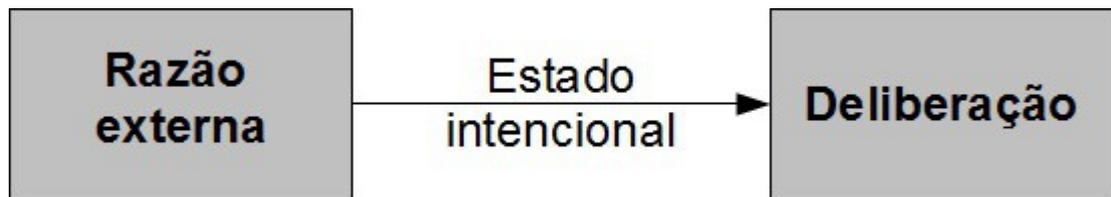
Vimos que a ação livre depende de os estados mentais poderem ser acerca do que lhes é externo. Entretanto se a ação tiver como causa suficiente uma crença ou um desejo – ou até ambos – não podemos dizer que ela é livre, visto que em tal caso não há o rompimento com a relação de causa e efeito. Para haver tal rompimento, como vimos, é necessário que o móbil da ação tenha origem em algo independente dos desejos. Ou seja, a origem da motivação da ação livre deve ser externa ao próprio estado mental, isto é, uma razão para a ação. A razão independente-do-desejo é um motivador para a ação livre. (SEARLE, 2013, p. 140). É a partir dela que surge a vontade de agir. Isto se dá no âmbito da intencionalidade, sendo a racionalidade um fator restritor dessa. A racionalidade reconhecadora (SEARLE, 2013, p. 127) – que é parte da intencionalidade – cria uma razão interna para a ação a partir de um motivador externo. Através deste processo da racionalidade reconhecadora, o motivador externo torna-se um motivador interno. (SEARLE, 2013, p. 197). Se este processo descrito for voluntário, não poderá ser passivo. Ora, raciocinar é um processo voluntário: não é experimentado como passivo. Além disso, raciocinar, sendo um processo, ocorre no tempo e exige uma unificação em um “eu”, ou seja, conforme Searle, uma entidade que experimenta tal processo não como um feixe passivo de percepções, mas como agente voluntário e, por isso, consciente.

Podemos diferenciar uma ação voluntária de uma involuntária. Segundo Searle (2013, p. 95-99), a ação involuntária não exige intencionalidade intrínseca nem consciência, portanto, em tal caso, o agente não necessariamente é um “eu”. Por outro lado, a ação voluntária pressupõe intencionalidade intrínseca e requer que o agente seja consciente. Assim, na ação voluntária, portanto racional, o agente é um “eu”. (SEARLE, 2013, p. 97). É a este “eu” que, por agir voluntariamente, pode ser imputada responsabilidade. Assim, temos dois tipos de agentes: o que pode ser responsabilizado e aquele ao qual não cabe responsabilidade.

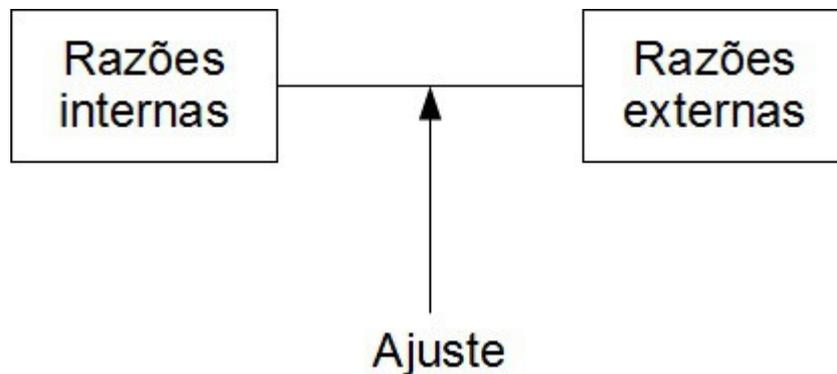
O segundo age por causa e efeito; o primeiro age com base em sua capacidade de intencionalidade intrínseca.

Para ficar mais claro, podemos esquematizar a posição de Searle (2013, p. 122-123) da seguinte forma.

**Figura 1 - Razão externa operando na deliberação**



**Figura 2 - Situação racional ideal**



O agente que delibera o faz somente por razões internas. Estas representam razões externas. Então as razões internas “procuram” ajustar-se às razões externas, as quais só operam na deliberação do agente porque este possui algum estado intencional que faz tal representação.

## 5.1 LIVRE-ARBÍTRIO, IA FORTE E RESPONSABILIDADE MORAL

Sendo a questão da intencionalidade intrínseca – como vimos – um ponto-chave para pesquisas acerca da questão do livre-arbítrio, propomos então verificar quais implicações ela promove na questão da possibilidade ou da impossibilidade da existência da inteligência artificial forte. A importância ética desta questão reside justamente em que seres capazes de possuir estados intencionais são passíveis de ser responsabilizados por seus atos. Se concluirmos que as máquinas atuais são capazes de intencionalidade intrínseca, em princípio, devemos lhes atribuir responsabilidade (moral ou legal).

Searle pensa que as máquinas atuais não sejam capazes de intencionalidade intrínseca por crer não serem elas conscientes. Isso se o que elas fazem é só uma manipulação formal de símbolos (SEARLE, 1998, p. 221-222) – os quais somente são símbolos para aquele que exerce a sua capacidade de intencionalidade intrínseca. Assim, como pretensão de resposta definitiva ao argumento do “jogo da imitação” de Turing, Searle (1980) apresentou, em seu artigo “Mentes, cérebros e programas”, o Argumento do Quarto Chinês. Ele iniciou o texto diferenciando a inteligência artificial (IA) no sentido forte daquela no sentido fraco. A IA no sentido fraco, para Searle (1980), no que se refere ao estudo da mente, apenas considera o computador como uma ferramenta. Já, de acordo com a IA no sentido forte, um computador adequadamente programado é uma mente, ou seja, consegue entender e tem estados cognitivos e, então, os programas explicariam a capacidade cognitiva humana. Dadas tais definições, Searle (1980) sustenta que para os defensores da IA no sentido forte, se computadores forem programados de forma adequada, literalmente terão estados cognitivos (SEARLE, 1980, p. 417).

O Argumento do Quarto Chinês apresentado no artigo supracitado é um experimento mental em que se propõe que um homem que compreende apenas o idioma inglês esteja trancado em um quarto. Tal homem recebe textos em chinês por uma fenda e deve, utilizando

um livro de regras, relacioná-los com caracteres, também em chinês, que devem ser repassados por outra fenda ao outro lado da sala. Para o homem no quarto, os caracteres em chinês não têm significado algum. Contudo para as pessoas de fora da sala, que entendem chinês, os textos que entram e saem do quarto são respectivamente perguntas e respostas. Tal argumento procurou mostrar como a semântica não é intrínseca à sintaxe e que um computador não compreende nada, pelo menos no sentido em que dizemos que um ser humano compreende.

O que Searle defende é que um programa instanciado não é condição suficiente para que haja compreensão, porque a manipulação de símbolos formais não tem intencionalidade intrínseca. Isto porque os símbolos só simbolizam algo nas mentes de quem insere o *input* e de quem interpreta o *output*. Logo, a intencionalidade somente existe na mente destes. Ele diz haver três pontos em que esbarram os defensores da relação de equivalência entre mente/cérebro e *software/hardware*, a saber (SEARLE, 1980): que um programa é diferente da sua realização (por isso o mesmo programa pode realizar-se de várias formas sem que haja intencionalidade intrínseca); que estados intencionais são definidos pelo seu conteúdo – não pela sua forma (o programa, por sua vez, é puramente formal); e que um programa não é produzido por um computador como a mente é produzida pelo cérebro.

O filósofo estadunidense crê que as pessoas confundem simulação com duplicação (SEARLE, 1980, p. 423), ao defender a relação de equivalência citada, por causa da tradição behaviorista, assim como também fazem confusão com a noção de informação, pois o fato de que quem insere o *input* e quem interpreta o *output* usem símbolos não significa que a máquina também o faça – a mesma só tem sintaxe e não semântica. Ademais a noção de informação pressupõe intencionalidade intrínseca porque nada é informação em si, toda informação o é para um “eu” que a reconhece como tal.

A posição searleana é a de que o cérebro produz causalmente a consciência, bem como a intencionalidade intrínseca e que os estados intencionais não são caracterizados por sua forma, mas por seu conteúdo. Como um programa é puramente formal, a sua instanciação não é suficiente para produzir a intencionalidade intrínseca (SEARLE, 1980) – “a sintaxe em si não é suficiente para o conteúdo semântico” (SEARLE, 2010, p. 17).

Concordamos com Searle quando ele defende em várias de suas obras que não é impossível criarmos artificialmente seres capazes de consciência e, conforme defendemos, de intencionalidade intrínseca. A condição para isto é que saibamos como os nossos cérebros produzem a consciência. O comportamento aparentemente inteligente de uma máquina atual é puramente "sintático", por assim dizer: é a extensão do que o programador (consciente e capaz de intencionalidade intrínseca) realizou. Quando jogamos videogame individualmente, não o fazemos contra a máquina, mas contra o programador. Ou seja, as máquinas atuais não são inteligentes: a inteligência está naqueles que as construíram e as programaram. Não sendo capazes de intencionalidade intrínseca, as máquinas atuais não possuem livre-arbítrio, apesar de muitas vezes aparentarem.

Tratamos anteriormente de que a ação voluntária pressupõe tanto a intencionalidade intrínseca quanto que o agente seja consciente. Como tais características não existem nas máquinas atuais, não podemos lhes atribuir nenhuma responsabilidade moral, somente a seus criadores, pois estes são capazes de intencionalidade intrínseca.

## **5.2 NEUROCIÊNCIA E LIVRE-ARBÍTRIO**

Uma objeção que poderia ser feita à afirmação de que somos capazes de ações livres por sermos capazes de intencionalidade intrínseca é sustentada nos experimentos de Benjamin Libet (1999). Ou seja, a objeção seria a de que tais experimentos comprovariam que não

somos capazes de ações livres. Se este fosse o caso, estaríamos no mesmo nível das máquinas atuais no que se refere ao livre-arbítrio e, da mesma forma, quanto à responsabilidade moral.

Libet abordou a questão do livre-arbítrio de forma experimental. Seus experimentos comprovaram que alterações elétricas no cérebro precedem em aproximadamente 350 a 400ms a intenção consciente do sujeito que age nas ações tidas como livres. Ele identificou portanto a existência de um potencial de disposição temporalmente anterior ao que Searle chama de intenção-na-ação. Ainda, segundo Libet (1999, p. 49), as alterações elétricas no cérebro – chamadas de potencial de disposição – precedem em aproximadamente 550ms a ativação do respectivo músculo, ou seja, o cérebro já está eletricamente configurado para um determinado movimento mais de meio segundo antes que ele ocorra. Ademais, como foi dito anteriormente, conforme Libet, quase meio segundo separa a alteração elétrica no cérebro da sensação do sujeito de estar decidindo agir.

Os experimentos de Libet partiram de alguns critérios para tentar concluir se possuímos livre-arbítrio ou não. Um deles era de que nada externo (sinais ou outras formas de controle) poderia influenciar o surgimento do ato tido como voluntário. Já o outro critério era justamente o de que o sujeito autor da ação deveria sentir que as suas escolhas eram livres, que ele poderia escolher agir de outra forma – que a sua ação estava sob o seu próprio controle. (LIBET, 1999, p. 47). Este último critério é o mesmo que Searle (2013) utiliza: sabemos que somos livres porque sentimos que poderíamos agir de outra maneira, que temos alternativas. Entretanto o próprio Searle (2013, p. 20) reconhece que isso não garante que somos livres: o livre-arbítrio poderia ser, mesmo assim, uma ilusão. Por outro lado, o próprio Libet (1999, p. 47) afirma que nós seríamos autômatos essencialmente sofisticados se a nossa subjetividade não tivesse poder causal, ou seja, se nossos sentimentos e intenções conscientes fossem epifenomenais.

O que conseguimos experimentalmente é algum conhecimento de como o livre-arbítrio pode operar. Mas não respondemos à questão de saber se os nossos atos de vontade consciente são totalmente determinados por leis naturais que regem as

atividades das células nervosas no cérebro, ou se os atos e as decisões conscientes para realizá-los podem avançar até certo ponto, independentemente do determinismo natural. A primeira destas opções tornaria o livre-arbítrio ilusório. O sentimento consciente de exercer a sua vontade seria então considerado como um epifenômeno, simplesmente um subproduto das atividades do cérebro, mas sem poderes causais próprios.(LIBET, 1999, p. 55).

O que o experimento de Libet mostrou é que antes que a ação livre aconteça “há uma preparação característica da atividade elétrica no cérebro” (CAVE, 2007, p. 555). Mas será que tal configuração elétrica prévia no cérebro (que Libet chama de potencial de disposição) seria causalmente suficiente para a consciência e a intencionalidade? Uma resposta positiva a esta pergunta significaria adesão ao determinismo, visto que a subjetividade estaria sujeita às leis deterministas do mundo físico. Entretanto o próprio Libet afirma que não é cientificamente comprovado que as funções e os eventos conscientes subjetivos são causados pela natureza determinista do mundo fisicamente observável. Para ele, tal suposição ainda é uma crença especulativa.(LIBET, 1999, p. 55-56). Sendo assim, a afirmação de Cave (2007, p. 555) de que a experiência consciente que temos de estar agindo no mundo – experiência que Searle identifica como sendo o livre-arbítrio – é só um tipo de acréscimo que não teria comprovação científica. A conclusão de Cave (2007, p. 555) sobre a epifenomenalidade do mental – baseada nos estudos de Libet – é portanto um tanto apressada. Pelo contrário, a epifenomenalidade do mental, reiteramos, é uma tese justamente incompatível com o conhecimento científico acerca da evolução das espécies. Se o epifenomenalismo – tese que assombra o fisicalismo – fosse o caso, o livre-arbítrio seria uma hipótese a ser descartada, o que reforça a tese de Searle de que esta não parece ser uma opção sensata.

Após os experimentos de Libet, outros neurocientistas fizeram revisões aprimoradas sobre os mesmos com equipamentos mais sofisticados, como descreve o artigo de Cinara Nahra (2013). Entretanto a autora critica a conclusão de que tais experimentos provam que o livre-arbítrio é uma ilusão. Um dos pontos de sua crítica relaciona-se com o que havíamos tratado anteriormente, ou seja, que na ação voluntária o agente é um “eu” e que é a este “eu”

que podemos atribuir responsabilidade moral porque age com base em sua capacidade de intencionalidade intrínseca. Neste sentido – partindo de que o agente é um “eu” – Cinara Nahra (2013, p. 185) acusa tais estudos (que corroborariam que o livre-arbítrio não existe) de pressuporem que o cérebro é uma entidade separada do “eu”. Para tal autora, esta pressuposição evidencia uma posição ingenuamente dualista. Isto porque ao defender que o nosso cérebro decide, por assim dizer, antes de tomarmos consciência da nossa decisão e que portanto não somos livres, determinadas revisões<sup>35</sup> dos experimentos de Libet dissociam o “eu” do corpo (ou do cérebro). Assim como Searle (2013, p. 159) afirma que os nossos estados mentais têm como base um *background* – uma rede de crenças, desejos e aptidões – Nahra (2013, p. 187) defende que a nossa decisão é influenciada por determinados mecanismos que são inconscientes. Ela afirma ainda que “o fato de que nós já teríamos escolhido algo antes de ter consciência disto não faz com que esta decisão não seja livre”. (NAHRA, 2013, p. 186). Esta afirmação é compatível com o *background* searleano visto como um “conjunto de estruturas cerebrais” (SEARLE, 2013, p. 62) que teria condições de nos deixar aptos a ativar o sistema de intencionalidade, sendo este imprescindível para a tomada de decisão livre. É importante notar que os experimentos citados por Nahra (2013, p. 187 – 188) não são tão precisos, pois ficaram apenas “10% acima do nível de acaso”. Ou seja, tais resultados demonstram mais uma tendência para agir do que uma decisão prévia. Com efeito, o *background* searleano não consiste de estados intencionais (SEARLE, 2013, p. 62), mas serve de base para a estrutura da racionalidade que, por sua vez, é interna à intencionalidade. Então o *background* também não determina a ação.

Nahra não pretende minimizar a importância da neurociência nos estudos sobre o livre-arbítrio, mas mostrar como tal ramo da ciência não prova que o livre-arbítrio não existe.

A filósofa afirma que os estudos neurocientíficos

---

<sup>35</sup> Nahra trata dos trabalhos de Chun Soon, John D. Haynes e Stefan Bode que, segundo ela, valem-se de técnicas mais avançadas que aquelas utilizadas por Libet e que pretensamente poderiam corroborar que o livre-arbítrio não existe.

Embora não mostrem que não temos livre arbítrio, mostram que há determinadas condições que limitam o exercício do nosso livre arbítrio, e fazem com que algumas de nossas ações não sejam voluntárias, consequentemente limitando nossa responsabilidade moral. [...] A neurociência, e particularmente a neurociência da ética, então, não ameaça a noção de responsabilidade moral e legal, ao contrário, ela pode ser usada como mais um elemento capaz de promover a justiça, evitando a que venhamos a responsabilizar aqueles que não podem ser responsabilizados, e com isso refinando a responsabilização moral e criminal na sociedade, de modo que venhamos a caminhar na direção de um mundo no qual a responsabilidade moral, a culpabilidade e também a punição sejam atribuídas apenas aqueles que a mereçam. (NAHRA, 2013, p. 197).

A neurociência não elimina portanto a responsabilização moral. Isto porque responsabilizar moralmente só se faz sobre um “eu”. Se tratarmos o cérebro como algo separado desse “eu”, cometeremos um equívoco. Ou seja, ao invés de o livre-arbítrio ser uma ilusão, esta reside, na verdade, em afirmar que “o livre-arbítrio existe independentemente do nosso cérebro” (NAHRA, 2013, p. 194) – o que seria uma postura dualista ingênua. A posição de Nahra portanto corrobora a tese searleana de que o livre-arbítrio não é uma ilusão. Podemos ainda assumir a seguinte interpretação acerca dos estudos neurocientíficos citados: de que o cérebro evidencia uma disposição prévia por escolher uma alternativa no lugar de outra, mas não elimina a importância da intencionalidade intrínseca na decisão.

Assim como Cinara Nahra não ignora os avanços da neurociência, a Searle – como é apontado por Cave (2007, p. 558) – não cabe a crítica de ignorar os avanços científicos. Searle leva em consideração as descobertas da ciência sem ignorar as questões sobre as quais os filósofos têm se debruçado ao longo de séculos, sendo a do livre-arbítrio uma das mais importantes, sobretudo por suas implicações no direito e na ética.

### **5.3 NATURALISMO BIOLÓGICO, MORAL EVOLUCIONISTA E UNIVERSALIDADE DAS LEIS MORAIS**

Já afirmamos que o naturalismo biológico de Searle é mais compatível – do que o fisicalismo e o funcionalismo – com a teoria de Darwin que trata da evolução das espécies. Dissemos que o livre-arbítrio, dentro desta abordagem, é uma vantagem evolutiva. Sendo assim, como a questão do livre-arbítrio está intimamente relacionada com a da

responsabilidade moral, como poderíamos relacionar moral com evolução? Poderíamos derivar um “deve” de um “é”? Em caso positivo, como ficaria a questão da universalidade das leis morais?

Segundo Chediak (2003, p. 47), uma das formas de relacionar moral com evolução é conceber que o biológico em interação com a cultura resultaria em moral. Mas uma objeção que poderia ser feita é a seguinte. O que o biológico fornece para o comportamento moral? Fornece algumas habilidades que são vantagens evolutivas, a saber, a de antecipação de consequências, a de distinguir pelo que é mais desejável e a de poder escolher entre alternativas dadas. (CHEDIAK, 2003, p. 48). Esta última inclusive está intimamente ligada à noção de livre-arbítrio.

Entretanto ainda se poderia objetar que tais habilidades não têm relação alguma com o cumprimento de promessas, com deveres ou obrigações. Tal objeção tem relação com a afirmação humeana – a Lei de Hume<sup>36</sup> – de que não se pode derivar um “deve” de um “é”, ou seja, com a suposta falácia naturalista (da qual trataremos adiante).

Por outro lado, se considerarmos a afirmação de que a intencionalidade intrínseca faz parte do mundo natural e que a ação livre está intimamente relacionada à intencionalidade, temos que discordar de Hume. Talvez não possamos derivar um “deve” de um “é” diretamente, mas há, de alguma maneira, uma relação. Searle (2013, p. 192) defende que há estruturas deônticas que compõem as estruturas intencionais as quais vinculam os membros de uma sociedade humana. O indivíduo inserido nesta sociedade tem o poder de criar livremente – e para si mesmo – razões independentes-do-desejo. Tal indivíduo, como agente

---

<sup>36</sup> “Em todo sistema de moral que até hoje encontrei, sempre notei que o autor segue durante algum tempo o modo comum de raciocinar, estabelecendo a existência de Deus, ou fazendo observações a respeito dos assuntos humanos, quando, de repente, surpreendo-me ao ver que, em vez das cópulas proposicionais usuais, como é e não é, não encontro uma só proposição que não esteja conectada a outra por um deve ou não deve. Essa mudança é imperceptível, porém da maior importância. Pois como esse deve ou não deve expressa uma nova relação ou afirmação, esta precisaria ser notada e explicada; ao mesmo tempo, seria preciso que se desse uma razão para algo que parece totalmente inconcebível, ou seja, como essa nova relação pode ser deduzida de outras inteiramente diferentes. [...] [A] distinção entre vício e virtude não está fundada meramente nas relações dos objetos, nem é percebida pela razão”. (HUME, *Tratado*, Livro III, Parte I, Seção I).

consciente que é, reconhece racionalmente os motivadores os quais derivam da assunção intencional e voluntária de obrigações, de cumprimento de promessas e deveres. Portanto a tese humeana de que é impossível para a razão ter o efeito de causar ou prevenir um desejo ou volição – a de que a razão é escrava das paixões (HUME, *Tratado*, Livro II, Parte III, Seção III) – parece ser falsa dado que vai contra a atribuição de um papel central à intencionalidade intrínseca na ação humana. Pelo que vimos anteriormente, posicionar-se contra o papel central da intencionalidade na ação humana e contra o seu caráter de vantagem evolutiva é negar o que conhecemos de biologia. Portanto, se a intencionalidade tem origem evolutiva, temos necessariamente que concluir que a moralidade também tem. Nesta mesma linha, Chediak (2003, p. 57) defende que

É bastante razoável considerar-se que a moralidade possa ter origem evolutiva, estando, por isso, assentada nos sentimentos que estabelecem atitudes de aprovação e de reprovação com relação às ações dos indivíduos, mas não há dúvida que as ações morais, por se expressarem em juízos, estão também sujeitas à crítica e à justificação. A justificação visa a fornecer razões a favor de determinadas normas, estabelecendo a razão da crença.

Neste sentido, Searle (2013, p. 190-191) defende que reconhecer uma razão para realizar uma ação – o que é um ato intencional – já é aceitá-la como uma razão para querer aceitar realizar tal ação. É desta maneira que as razões independentes-do-desejo motivam o agente. De fato, se as leis morais forem redutíveis simplesmente às suas origens evolutivas, não serão universais. A universalidade das leis morais parece estar no reconhecimento das razões. Tal reconhecimento é um ato intencional.

#### **5.4 A FALÁCIA NATURALISTA**

A falácia naturalista, já citada anteriormente, tem sua formulação original atribuída a Hume (*Tratado*, Livro III, Parte I, Seção I) e consistiria em extrair conclusões valorativas a partir de premissas que não possuem conteúdo valorativo, apenas fatos. Em outras palavras,

supostamente não se poderia derivar um “deve” de um “é”, pois haveria uma descontinuidade, um hiato, entre fatos e valores. Entretanto, alguns filósofos – sobretudo os naturalistas –, dentre os quais podemos incluir o próprio Searle, não consideram que haja tal falácia. Ademais, se aceitarmos que não podemos derivar um “deve” de um “é”, não poderemos aceitar que a moralidade tenha origem evolutiva – o que vai contra o que concluímos anteriormente. Precisamos, portanto, verificar se a falácia naturalista se sustenta.

Para Karla Chediak (2003, p. 58) o evolucionismo não pretende nos dizer como agir mas deixar claro o que a evolução biológica tem a ver com a evolução cultural. Mas Searle (1964) vai mais além. Ele ataca a chamada “falácia naturalista”. Afirma que se pode sim derivar um “deve” de um “é”. Ou melhor, Searle (1964, p. 55) defende que o modelo clássico não é bem-sucedido porque confunde declarações que são acerca de fatos brutos com aquelas que são sobre fatos institucionais. Tal erro viria de outro, ou seja, o de não fazer a distinção entre regras regulativas e regras constitutivas. Vamos analisar melhor. O que são, para Searle, regras regulativas e constitutivas? Temos, em suas próprias palavras a seguinte explicação: “regras regulativas regulam atividades cuja existência é independente das regras; regras constitutivas constituem (e também regulam) formas de atividade cuja existência é logicamente dependente das regras”. (SEARLE, 1964, p. 55). Um exemplo dado por ele provém do jogo de xadrez: as regras do jogo são constitutivas do mesmo. Não há como ter o jogo sem as suas regras. As regras do jogo não são fatos brutos: são fatos institucionais (utilizando o termo aplicado por Searle). Mas tais fatos institucionais são intencionais: dependem da intencionalidade daqueles que jogam.

Ora, alguns sistemas de regras constitutivas – tais como as leis de trânsito – “envolvem obrigações, compromissos e responsabilidades” (SEARLE, 1964, p. 57). O modelo clássico, para Searle, trata os fatos institucionais (intencionais) como se fossem fatos brutos. Estes últimos não constituem sistemas de regras intencionais. Entretanto obrigações,

compromissos e responsabilidades estão relacionados a afirmações: uma promessa “cria” uma obrigação por convenção, pois um ato de fala não deixa de ser um ato convencional (SEARLE, 1964, p. 50), ou seja, possui intencionalidade. A forma lógica das promessas seria a seguinte: “Se C então (se U então P)”<sup>37</sup> (SEARLE, 1964, p. 45). Assim, para Searle, desde que as condições permaneçam as mesmas, nós podemos derivar um “deve” de um “é” porque não há premissas com valor moral nesta forma lógica e, portanto, "premissas factuais podem implicar conclusões valorativas"(SEARLE, 1964, p. 58). Alguém poderia objetar o seguinte. Qual seria o problema de termos premissas com valor moral na forma lógica acima? Se tivéssemos premissas com valor moral, um “deve” seria derivado de outro “deve” – o que seria um regresso ao infinito. Em suma, o que Searle defende é que a linguagem, por si só, institui o dever. Podemos ainda acrescentar que se a linguagem se dá sobre o que é externo ao próprio estado mental, temos que ela depende de intencionalidade intrínseca. Por conseguinte temos que o dever tem origem na intencionalidade.

Searle supõe então ter contornado a falácia naturalista. Considerando ainda que a motivação para a ação reside na racionalidade reconhecedora, ou melhor, que o reconhecimento de uma razão para agir é suficiente para motivar a ação (SEARLE, 2013, p. 196-205), ou seja, que a razão interna para agir vem do reconhecimento de um motivador externo, temos que a moralidade tem relação estreita com a intencionalidade que, por sua vez, está relacionada com o livre-arbítrio.

---

<sup>37</sup> C seriam as condições, U as afirmações e P as promessas.

## 6 CONSIDERAÇÕES FINAIS

Procuramos, ao longo desta pesquisa, analisar se o naturalismo biológico de John Searle, comparado ao fisicalismo e ao funcionalismo, oferece uma maior possibilidade de compatibilização do livre-arbítrio com o princípio da causalidade. Em outras palavras, buscamos averiguar se o naturalismo biológico searleano apresenta-se como uma alternativa mais viável que as correntes fisicalistas e funcionalistas nas pesquisas que tratam da questão do livre-arbítrio. Para tanto, tomamos como estratégia de pesquisa identificar os pressupostos destas linhas de pensamento da filosofia da mente, bem como as consequências filosóficas da adesão a cada uma delas. Entretanto, para seguir tal caminho, buscamos um conceito que nos servisse de fio condutor, que pudesse ser o núcleo da análise comparativa objetivada. Como verificamos que a ação livre depende de que os estados mentais daquele que age possam ser sobre o que é externo aos próprios estados mentais, a questão da intencionalidade intrínseca tomou o papel de núcleo da pesquisa. Tal entendimento mostrou-se um avanço que, articulado ao diagnóstico acerca das teses do naturalismo biológico, do fisicalismo e do funcionalismo, contribuiu para indicar que a primeira corrente é mais compatível com o princípio da causalidade e o livre-arbítrio. Ademais, a investigação demonstrou que a filosofia da mente de John Searle associa-se à teoria darwinista da evolução com maior propriedade. Ou seja, o naturalismo biológico não vai contra o conhecimento contemporâneo de biologia.

Vimos que fisicalistas e funcionalistas aceitam – em geral – a categorização entre físico e mental e que isso os faz se depararem com certas dificuldades perante a questão do livre-arbítrio já quando tratam da ontologia da consciência. Os primeiros acabam por reduzir a consciência e toda subjetividade a eventos físicos ou, por outro lado, terminam por simplesmente eliminar tais aspectos da realidade. Respectivamente, as consequências são a adesão ao determinismo (consequentemente ao epifenomenalismo) e ao eliminativismo. Ora, é inaceitável a postura eliminativista, como vimos, porque nos é evidente que somos

conscientes e, se assim somos, a consciência deve – como defende Searle – fazer parte do mundo. Por outro lado, se o mental – já utilizando as categorias rechaçadas por Searle – fosse epifenomenal ou superveniente, a sua vantagem evolutiva seria nula. Ademais, ao longo desta pesquisa, um ponto que se tornou, de certa forma, um dos pilares da argumentação foi que o livre-arbítrio é uma vantagem evolutiva. Ora, sabendo que somos conscientes e considerando a teoria evolucionista de Darwin, somos forçados a tratar a consciência como uma vantagem evolutiva. Neste sentido, as sensações de decidir e de agir, enfim, de atuar sobre o mundo, têm de ser, por sua vez, vantagens evolutivas. Outrossim, a racionalidade como parte da intencionalidade, também tem sua função evolutiva e, se não tivesse poder causal, as crenças e os desejos seriam condições causalmente suficientes para a ação – o que tornaria o livre-arbítrio impossível. É importante lembrarmos que o naturalismo biológico searleano defende que a consciência é parte da realidade, sendo também uma vantagem evolutiva. Portanto, um dos progressos obtidos na análise comparativa do naturalismo biológico com o fisicalismo foi o de que o primeiro consegue articular com mais propriedade a sua ontologia da consciência com a teoria evolucionista de Darwin. Na mesma linha, a questão da intencionalidade intrínseca como requisito para uma ação independente-do-desejo mostrou-se um ponto forte da posição searleana. Por tais aspectos, vimos que o naturalismo biológico de John Searle, comparado ao fisicalismo, apresenta-se como um caminho mais viável para uma pesquisa acerca do livre-arbítrio.

Notadamente avaliamos que a intencionalidade intrínseca é um requisito para a ação livre, isto é, o livre-arbítrio pressupõe intencionalidade intrínseca. Da mesma forma, a intencionalidade intrínseca somente se dá quando há subjetividade. Sendo assim ficou claro nesta pesquisa que a subjetividade é ontologicamente anterior ao livre-arbítrio. Tal progresso serviu para que pudéssemos analisar comparativamente também o naturalismo biológico com o funcionalismo. Abordamos o funcionalismo levando em conta que tal corrente padece das

mesmas dificuldades que o fisicalismo tanto na questão da ontologia da consciência quanto na do livre-arbítrio, pois as teses funcionalistas adotam os mesmos pressupostos que as fisicalistas, tais como as categorias tradicionais e a negação da intencionalidade intrínseca. A diferença desta variante do fisicalismo mostrou-se ser o tratamento da mente como sendo uma função, ou seja, o cerne da abordagem funcionalista é o aspecto formal, tentando identificar a sintaxe com a intencionalidade. Entretanto, mostramos que a sintaxe – além de não ser o mesmo – não é suficiente para a intencionalidade intrínseca. Ademais, a posição funcionalista, sobretudo a dennettiana, admite somente a intencionalidade derivada, o que, por conseguinte, implica não reconhecer um ponto de vista de primeira pessoa e, por isso, não admitir a existência de um “eu” – presumido em qualquer ação. Ainda mais, o funcionalismo não consegue escapar de um dilema entre identificar o “eu” com função ou simplesmente negar a sua existência. Os dois caminhos se nos apresentaram como problemáticos, visto que assumir a existência do “eu” e considerar o cérebro como algo dissociado dele, implica um dualismo de substâncias. Por outro lado, observamos que negar a existência do “eu” acarreta a negação da intencionalidade intrínseca, posto que o ponto de vista de primeira pessoa é um pressuposto. Isto, articulado à nossa conclusão de que a intencionalidade intrínseca é um requisito necessário para o livre-arbítrio, nos levou a entender que o funcionalismo, assim concebido, enfrenta sérias dificuldades ao tratar da questão do livre-arbítrio. Destarte demonstramos que, entre o naturalismo biológico e o funcionalismo, o primeiro apresenta-se como uma alternativa mais interessante para as pesquisas que tratam da questão do livre-arbítrio.

Apesar desta vantagem, por assim dizer, uma crítica que poderia ser feita ao naturalismo biológico searleano é a de que ele não cumpre com tudo o que promete, sobretudo ao tentar explicar como o livre-arbítrio é possível. Tal crítica poderia fundamentar-se no fato de Searle estabelecer um limite para a explicação da ação livre, ou seja, quando

postula o *gap*, ele deixa um hiato em parte da sua explanação. Ainda, não obstante os argumentos apresentados por Searle, parece faltar uma explicação do estatuto ontológico da intenção-na-ação. Neste mesmo sentido, poderíamos agregar outra objeção: a de que talvez lhe falte uma explicação para a intenção-na-ação não possuir causas suficientes e ao mesmo tempo – no caso da ação livre – ser causalmente suficiente para a ação. Entretanto, o autor sempre reconheceu que o limite das suas teses se encontra no nosso desconhecimento acerca de como o cérebro produz a consciência. Além do mais, não se deve esperar uma explicação definitiva para a questão do livre-arbítrio, mas uma abordagem que não esbarre em contradições nem ignore o conhecimento científico. Desta forma, o naturalismo biológico searleano consegue esclarecer mais do que as outras correntes filosóficas, com as quais comparamos, acerca de como pode a ação livre ter a origem da sua motivação no que é externo ao estado mental que a faz ser realizada. Ou melhor, Searle consegue aprofundar a questão do livre-arbítrio explicando como uma razão, independente-do-desejo, pode ser a origem para a ação livre; isto sem deixar de lado os experimentos feitos no campo da neurociência.

Como dissemos, Searle leva os experimentos neurocientíficos em consideração na sua abordagem da questão do livre-arbítrio. Porém, a sua postulação de um “eu” vai contra interpretações de tais experimentos que procuram dissociar o cérebro desse “eu”, como se o primeiro fosse algo à parte. Neste sentido, foram valiosos os argumentos de Cinara Nahra ao defender que a neurociência não elimina o livre-arbítrio. Com efeito, considerar que a mente pode ser vista como algo separado do cérebro tem as consequências, já observadas, do dualismo e do epifenomenalismo. O cérebro é, assim, parte constitutiva do “eu”.

Ora, podemos articular a questão da inteligência artificial forte com a questão do “eu”. Isto porque Searle considera que o cérebro produz causalmente a consciência. Por outro lado, programas de computador são puramente formais: podem ser implementados em meios

físicos particularmente distintos. Por serem somente formais, não são capazes de intencionalidade intrínseca – o que implica a impossibilidade de programas de computador serem identificados com um “eu”. Entretanto vimos que não faz sentido atribuir responsabilidade moral senão a um “eu”. Isto porque somente o “eu” que age voluntariamente é passível de ser responsabilizado moralmente. Daí a importância das conclusões acerca da intencionalidade intrínseca articuladas com as questões do livre-arbítrio e da inteligência artificial forte: elas nos ajudam a esclarecer a delimitação de seres aos quais a responsabilidade moral pode ser atribuída.

Porém, vimos que se há uma restrição de seres aos quais podemos atribuir responsabilidade moral, temos de aceitar que podemos extrair um “deve” de um “é”. Isto porque somente seres capazes de intencionalidade intrínseca são passíveis de responsabilização. Por outro lado, se a intencionalidade tem origem evolutiva, a moralidade também tem. Mas se não pudéssemos derivar um “deve” de um “é”, não haveria como afirmarmos a origem evolutiva da moralidade – o que, considerando as conclusões às quais chegamos acerca da intencionalidade, não nos é aceitável.

Podemos fazer ainda o seguinte balanço das conclusões anteriores: se, como vimos (após a análise da forma lógica sem premissas valorativas), a linguagem – que cria o dever – depende da capacidade de intencionalidade intrínseca, temos que esta última é a origem do dever. Por isso, afirmamos que a moralidade depende da intencionalidade a qual fundamenta o livre-arbítrio. Mas também analisamos que não poderíamos concluir pela universalidade das leis morais apenas por sua origem evolutiva. Ficaríamos sem uma explicação de como as leis morais poderiam ter um caráter universal. Então, como a razão interna para agir surge do reconhecimento de um motivador externo, pela tese searleana, e como tal reconhecimento é um ato intencional, a universalidade das leis morais residiria justamente no reconhecimento de tal razão. É importante salientar que, neste ponto, não procuramos fornecer uma prova da

universalidade das leis morais – o que seria algo deveras pretensioso –, mas antes apresentar uma abordagem que contribua com premissas a partir das quais um caminho para pesquisas em ética seja esboçado.

Esta pesquisa obviamente não tem a pretensão de colocar um ponto final à discussão acerca do livre-arbítrio, entretanto uma tese tornou-se mais sólida ao longo dela, ou seja, que a ação livre pressupõe intencionalidade intrínseca e consciência. A veracidade da mesma nos indica que o naturalismo biológico tem a vantagem de ser mais compatível com a teoria darwinista – quanto ao livre-arbítrio – do que as correntes da Filosofia da Mente com as quais a comparamos. Ademais, se o livre-arbítrio depende de intencionalidade intrínseca, a inteligência artificial forte se nos mostrou não ser possível a não ser que possamos criar artificialmente seres conscientes a partir do conhecimento (que ainda não temos) de como o nosso cérebro produz a consciência.

## REFERÊNCIAS

- ABRANTES, Paulo C. C. T. Nagel e os limites de um reducionismo fisicalista. **Cad. Hist. Fil. Ci.**, Campinas, Série 3, v. 15, n. 1, pp. 223-244, jan.-jun. 2005.
- ARISTÓTELES. **Ética a Nicômaco**. 1.ed. Edipro: Bauru, 2002.
- BENSUSAN, Hilan. Como levar estados mentais a sério (epifenômenos e fingimentos). **Manuscrito**, Campinas, v. 25, pp. 79-88, mar. 2002.
- BIZARRO, Sara F. Navalha de Ockham. In: BRANQUINHO, J.; MURCHO, D.; GOMES, N. G. **Enciclopédia de termos lógico-filosóficos**. São Paulo: Martins Fontes, 2006. pp. 535-536.
- BLACKBURN, Simon. **Dicionário Oxford de Filosofia**. Rio de Janeiro: Jorge Zahar Editor, 1997.
- BOEHNER, Philotheus; GILSON, Etienne. **História da Filosofia Cristã**. 8.ed. Petrópolis: Vozes, 2003.
- BRITO, Evandro O. Franz Brentano e a descrição dos atos psíquicos intencionais: uma exposição esquemática do manuscrito *Psychognosie* de 1890. **Ágora Filosófica**, Recife, v. 1, n. 1, pp. 87-114, jan.-jun. 2012. Disponível em: <<http://www.unicap.br/ojs-2.3.4/index.php/agora/article/download/164/150>>. Acesso em: 9 mai. 2014.
- CAVE, Stephen. [Resenha de] Eu penso, portanto sou, eu penso. **Revista Latinoamericana de Psicopatologia Fundamental**, v. 10, n. 3, p. 554-558, set. 2007. Disponível em: <<http://www.redalyc.org/articulo.oa?id=233017590014>>. Acesso em: 12 ago. 2014.
- CESCON, Everaldo. **Fenomenologia da consciência e da mente**. Caxias do Sul: EDUCS, 2013.
- CHALMERS, David J. **The conscious mind: in search of a fundamental theory**. Nova Iorque: Oxford University Press, 1996.
- CHEDIAK, Karla. Notas sobre a concepção evolucionista da moral. **Episteme**, Porto Alegre, n. 16, p. 45-59, jan.-jun. 2003.
- CHURCHLAND, Paul M. **Matéria e consciência: uma introdução contemporânea à filosofia da mente**. São Paulo: Editora UNESP, 2004.
- COSTA, Claudio. Livre Arbítrio: como ser um bom compatibilista. **Princípios: Revista de Filosofia**, Natal, v. 7, n. 8, p. 19-33, jan.-dez. 2000. Disponível em: <<http://dialnet.unirioja.es/descarga/articulo/2564885.pdf>>. Acesso em: 28 out. 2014.
- \_\_\_\_\_. **Filosofia da mente**. Rio de Janeiro: Jorge Zahar Ed., 2005.
- DAVIDSON, Donald. **Essays on actions and events**. 2. ed. Oxford: Oxford University Press, 2001.

DEL AMO, José Antonio Guerrero. Las dificultades del compatibilismo de Dennett. **Thémata: Revista de filosofía**, n. 39, p. 97-104, 2007. Disponível em: <<http://dialnet.unirioja.es/servlet/articulo?codigo=3310306>>. Acesso em: 1 jun. 2014.

DENNETT, Daniel C. **Consciousness explained**. Nova Iorque: Little, Brown and Company, 1991.

\_\_\_\_\_. **La libertad de acción: un análisis de la exigencia del libre albedrío**. Barcelona: Gedisa, 1992.

\_\_\_\_\_. **Tipos de mentes: rumo a uma compreensão da consciência**. Rio de Janeiro: Rocco, 1997.

\_\_\_\_\_. **Brainstorms: ensaios filosóficos sobre a mente e a psicologia**. São Paulo: Editora UNESP, 2006.

FLANAGAN, Owen J. **The science of the mind**. 2. ed. Cambridge: MIT Press, 1991.

FREGE, Gottlob. O Pensamento – Uma Investigação Lógica. **Cad. Hist. Fil. Ci.**, Campinas, série 3, v. 8, n. 1, p. 177-208, jan.-jun. 1998.

GULICK, Robert Van. Functionalism. In: MCLAUGHLIN, Brian P.; BECKERMANN, Ansgar; WALTER, Sven. **The Oxford Handbook of Philosophy of Mind**. Oxford: Clarendon Press, 2011. Cap. 7, p. 128-151.

HODGES, Andrew. **Turing: um filósofo da natureza**. São Paulo: Editora UNESP, 2001.

HUME, David. **Tratado da natureza humana: uma tentativa de introduzir o método experimental de raciocínio nos assuntos morais**. São Paulo: Editora UNESP, 2000.

KIM, Jaegwon. Supervenience, determination, and reduction. **The Journal of Philosophy**, v. 82, n. 11, pp. 616-618., nov. 1985.

\_\_\_\_\_. **Mind in a physical world: an essay on the mind-body problem and mental causation**. Cambridge: MIT Press, 1998.

\_\_\_\_\_. **Physicalism, or something near enough**. New Jersey: Princeton University Press, 2008.

\_\_\_\_\_. Mental causation. In: MCLAUGHLIN, Brian P.; BECKERMANN, Ansgar; WALTER, Sven. **The Oxford Handbook of Philosophy of Mind**. Oxford: Clarendon Press, 2011. Cap. 1, p. 29-52.

KLAUDAT, André. Materialismo, causalidade e eventos no monismo anômalo. **Episteme**, Porto Alegre, n. 9, p. 47-59, jul.-dez. 1999.

LIBET, Benjamin. Do we have free will?. **Journal of Consciousness Studies**, v. 6, n. 8-9, p. 47-57, jan.-jun. 1999.

LOPES, Gustavo Cadar. **Determinismo causal e livre-arbítrio: o neocompatibilismo de Dennett e Frankfurt**. 2009. 90 f. Dissertação (Mestrado em Filosofia Moderna e Contemporânea) – Universidade Estadual do Rio de Janeiro, Programa de Pós-Graduação em

Filosofia Moderna e Contemporânea , Rio de Janeiro, 2009. Disponível em: <[http://www.livrosgratis.com.br/arquivos\\_livros/cp110344.pdf](http://www.livrosgratis.com.br/arquivos_livros/cp110344.pdf)>. Acesso em: 1 jun. 2014.

MARGONI, Lucas F. **O funcionalismo na filosofia da mente**. Porto Alegre: Editora Fi, 2013.

MASLIN, K. T. **Introdução à Filosofia da Mente**. São Paulo: Artmed, 2009.

MELNYK, Andrew. Physicalism. In: STICH, S. P.; WARFIELD, T. A. **The Blackwell Guide to Philosophy of Mind**. Cornwall: Blackwell Publishing, 2003. pp. 65-84.

NAHRA, Cinara. Neuroética, livre arbítrio e responsabilidade moral: a neurociência não prova que o livre-arbítrio é uma ilusão. **Dissertatio**, Pelotas, n. 38, pp. 181-199, 2013. Disponível em: <<http://www2.ufpel.edu.br/isp/dissertatio/revistas/38/9.pdf>>. Acesso em: 8 out. 2014.

PIRES, Jesuino Junior. Considerações sobre o conceito de intencionalidade em Edmund Husserl. **Kínesis**, v. 4, n. 7, p. 286-302, 2012.

PRATA, Tárík de Athayde. É incoerente a concepção de Searle sobre a consciência?. **Manuscrito**: Revista Internacional de Filosofia, v. 34, n. 2, p. 557-578, 2011.

\_\_\_\_\_. O modelo ontológico estratificado no naturalismo biológico de John Searle: uma controvérsia com Jaegwon Kim. **Discusiones Filosóficas**, v. 13, n. 21, p. 119-137, jul.-dez. 2012.

\_\_\_\_\_. É o naturalismo biológico uma concepção fisicalista?. **Principia**: an international journal of epistemology, v. 16, n. 2, p. 255-276, 2013.

PUTNAM, Hilary. **A triplacorda: mente, corpo e mundo**. Lisboa: Instituto Piaget, 2002.

RUSSELL, Bertrand. **A análise da mente**. Rio de Janeiro: Zahar Editores, 1976.

\_\_\_\_\_. **Ensaios céticos**. L&PM Editores: Porto Alegre, 2010.

SEARLE, John R. How to derive “ought” from “is”. **The Philosophical Review**, v. 73, n. 1, p. 43-58, jan. 1964. Disponível em: <<http://www.jstor.org/stable/2183201>>. Acesso em: 4 ago. 2014.

\_\_\_\_\_. Minds, brains, and programs. **Behavioral and brain sciences**, v. 3, n. 3, p. 417-424, 1980.

\_\_\_\_\_. **Mente, cérebro e ciência**. Lisboa: Edições 70, 1984.

\_\_\_\_\_. **O mistério da consciência**. São Paulo: Paz e Terra, 1998.

\_\_\_\_\_. **Rationality in action**. Cambridge: MIT Press, 2001.

\_\_\_\_\_. **Intencionalidade**. São Paulo: Martins Fontes, 2002.

\_\_\_\_\_. **A redescoberta da mente**. São Paulo: Martins Fontes, 2006.

\_\_\_\_\_. **Liberdade e neurobiologia**. São Paulo: Editora UNESP, 2007.

\_\_\_\_\_. **Consciência e linguagem**. São Paulo: Martins Fontes, 2010.

\_\_\_\_\_. **Razones para actuar**: una teoría del libre albedrío. Oviedo: Nobel, 2013.

SPARANO, Maria Cristina de Távora. **Linguagem e significado**: o projeto filosófico de Donald Davidson. Porto Alegre: EDIPUCRS, 2003.

TEIXEIRA, João de F. **Mentes e máquinas**: uma introdução à ciência cognitiva. Porto Alegre: Artes Médicas, 1998.

\_\_\_\_\_. **Como ler a filosofia da mente**. São Paulo: Paulus, 2008.

TRIPICCHIO, Adalberto; TRIPICCHIO, Ana C. **Teorias da mente**. Ribeirão Preto: Tecmedd, 2004.

TURING, Alan M. Computing machinery and intelligence. **Mind**, Oxford, New Series, v. 59, n. 236, pp. 433-460, out. 1950. Disponível em: <<http://mind.oxfordjournals.org/content/LIX/236/433.full.pdf>>. Acesso em: 9 mai. 2014.

WAKEFIELD, Jerome; DREYFUS, Hubert. Intentionality and the phenomenology of action. In: LEPORE, Ernest; VAN GULICK, Robert (Org.). **John Searle and his critics**. Oxford: Basil Blackwell, 1991. p. 259-270.

ZILIO, Diego. Fisicalismo na filosofia da mente: definição, estratégias e problemas. **Ciências e Cognição/Science and Cognition**, v. 15, n. 1, p. 217-240, 2010. Disponível em: <<http://www.cienciasecognicao.org/revista/index.php/cec/article/view/264/173>>. Acesso em: 9 mai. 2014.