



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**INSTITUTO DE CULTURA E ARTE**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA**

**SILMARA NATIFATE DA SILVA LIMA**

**A TEORIA DA CONSCIÊNCIA DE SEARLE E SUA CRÍTICA AO  
COGNITIVISMO**

**FORTALEZA**

**2023**

SILMARA NATIFATE DA SILVA LIMA

A TEORIA DA CONSCIÊNCIA DE SEARLE E SUA CRÍTICA AO COGNITIVISMO

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Mestre em Filosofia. Área de concentração: Filosofia da mente.

Orientador: Prof. Dr. Cícero Antônio Cavalcante Barroso.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- L71t Lima, Silmara Natifate da Silva.  
A teoria da consciência de Searle e sua crítica ao Cognitivismo / Silmara Natifate da Silva  
Lima. – 2023.  
85 f.
- Dissertação (mestrado) – Universidade Federal do Ceará, Instituto de cultura e Arte,  
Programa de Pós-Graduação em Filosofia, Fortaleza, 2023.  
Orientação: Prof. Dr. Cícero Antônio Cavalcante Barroso.
1. John Searle. 2. Consciência. 3. Subjetividade. 4. Cognitivismo. I. Título.

CDD 100

---

SILMARA NATIFATE DA SILVA LIMA

A TEORIA DA CONSCIÊNCIA DE SEARLE E SUA CRÍTICA AO COGNITIVISMO

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Mestre em Filosofia. Área de concentração: Filosofia da mente.

Aprovada em: 17/07/2023.

BANCA EXAMINADORA

---

Prof. Dr. Cícero Antônio Cavalcante Barroso (orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Luís Filipe Estevinha Lourenço Rodrigues  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Ralph Leal Heck  
Universidade Federal do Ceará (UFC)

Aos meus pais, Fátima Dias da Silva e Raimundo  
Nonato Lima.

## **AGRADECIMENTOS**

A minha família pelo contínuo apoio e amor do qual serei grata pela eternidade.

Ao meu orientador prof. Dr. Cícero Antônio Cavalcante Barroso, pela sua paciência, atenção e dedicação. Sem sua ajuda este trabalho não seria possível.

Aos professores Ralph Leal Heck e Luís Filipe Estevinha Lourenço Rodrigues pela disponibilidade e preciosas correções. Meu muito obrigada.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) pelo apoio financeiro a esta pesquisa (Código de Financiamento 001).

## RESUMO

Com a obra *The Rediscovery of the Mind* de 1992, o filósofo estadunidense John Searle apresentou ao mundo sua teoria geral da mente que tinha como proposta central tornar os fenômenos qualitativos e subjetivos em temas centrais nas discussões sobre a mente, dando-lhes um enfoque naturalista em contraposição ao enfoque dualista que ele via em boa parte das teorias contemporâneas. As pretensões da teoria do autor incluíam também: superar as teorias tradicionais no estudo da mente (mais especificamente, o dualismo e o materialismo); demonstrar que a consciência é o fenômeno mais fundamental das mentes; e refutar o Computacionalismo, em especial, a vertente chamada por ele de “*Cognitivismo*”. Este trabalho é sobre as duas últimas pretensões, isto é, visa apresentar a hipótese Searleana de que a consciência é central para caracterizar a mente, e contemplar, a partir disso, as razões pelas quais o autor reprova as ideias computacionalistas e rejeita completamente o cognitivismo. Por fim, deveremos refletir se a hipótese de centralidade da consciência se sustenta e se as críticas de Searle ao cognitivismo são justificadas.

**Palavras-chave:** John Searle; consciência; subjetividade; cognitivismo.

## ABSTRACT

In *The Rediscovery of the Mind*, the American philosopher John Searle presented to the world his general theory of the mind, which had as its central proposal to make qualitative and subjective specifications central themes in discussions about the mind, giving them a naturalistic approach in contrast to the dualist approach that he saw in most contemporary theories. The aims of the author's theory also included: overcoming traditional theories in the study of the mind (more specifically, dualism and materialism); demonstrate that consciousness is the most fundamental characteristics of minds; and refutes Computationalism, in particular, the aspect he called "*Cognitivism*". This work is about the last two claims, that is, to present the Searlean hypothesis that consciousness is central and fundamental to characterize the mind, and to contemplate, from this, the reasons why the author disapproves of computationalist ideas and rejects cognitivism. Finally, we must reflect on whether his hypotheses regarding the centrality of consciousness are supported and whether his criticisms of cognitivism are justified.

**Keywords:** John Searle; consciousness; subjectivity; cognitivism.



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>11</b>
<b>2</b>	<b>O PROBLEMA DA SUBJETIVIDADE</b> .....	<b>15</b>
2.1	A objetividade científica.....	15
2.2	A metodologia reducionista.....	19
2.2.1	<i>O problema do dualismo e o reducionismo na Filosofia da Mente</i> .....	22
2.2.2	<i>O reducionismo na Psicologia e o aspecto comportamental</i> .....	27
2.3	A Metáfora Computacional .....	31
2.3.1	<i>Mentes simbólicas e a Inteligência Artificial</i> .....	31
2.3.2	<i>O Funcionalismo</i> .....	35
2.3.3	<i>A Ciência Cognitiva e o surgimento do Cognitivismo</i> .....	36
2.4	Uma nova concepção de mente? .....	41
<b>3</b>	<b>A TEORIA SEARLIANA DA CONSCIÊNCIA</b> .....	<b>44</b>
3.1	Por que a consciência é irreduzível? .....	44
3.2	Consciência: subjetividade, qualitatividade e unidade .....	49
3.3	Intencionalidade .....	51
3.4	O Campo da Consciência e seus aspectos .....	54
3.5	A Inconsciência e o Princípio da Conexão.....	56
3.6	E agora? O que fazer com a consciência?.....	62
<b>4</b>	<b>AS CRÍTICAS DE SEARLE AO COGNITIVISMO E AS OBJEÇÕES QUE ELAS ENFRENTAM</b> .....	<b>65</b>
4.1	As limitações da Metáfora Computacional .....	65
4.1.1	<i>O Quarto Chinês</i> .....	67
4.1.2	<i>A Crítica da Razão Cognitivista</i> .....	70
4.2	Existem problemas nas alegações de Searle contra o Cognitivismo? .....	75
4.3	O problema da Forma Aspectual.....	78

<b>5</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>82</b>
	<b>REFERÊNCIAS .....</b>	<b>85</b>

## 1 INTRODUÇÃO

O que é a voz que escutamos em nossas mentes? São os nossos pensamentos? O que exatamente está causando isso? As perguntas que envolvem a mente são sem dúvida as mais intrigantes, afinal, ela faz parte do grupo de fenômenos que o homem ainda não compreende por completo, o que, sem dúvidas, torna as discussões sobre a mente um campo de investigação rico em hipóteses e multiplicidades teóricas. Na filosofia, os fenômenos mentais, como por exemplo, as crenças e o pensamento são temas contemplados por uma vasta tradição de teorias que vão desde os tempos da Grécia antiga à contemporaneidade.

Entre todos os tópicos de discussão relativos ao mental, o filósofo americano John Searle diz que a consciência é a temática mais significativa, pois ela carrega um problema fundamental: a condição subjetiva da consciência a torna um objeto de estudo inviável para as ciências ou qualquer outro campo de investigação que siga o princípio de objetividade científica. Searle acredita que a dificuldade das discussões sobre a consciência reside justamente na crença geral de que existe um conflito entre sua natureza subjetiva e a objetividade científica. O autor também considera que isso não é um problema exclusivo das ciências, pois na psicologia e na filosofia essa visão parece não se alterar. Ao menos não na grande maioria das teorias da mente.

Segundo Searle, o problema dessa situação é que ela não favorece aqueles que desejam compreender a real natureza das mentes, o que inclui, tanto a consciência como os aspectos subjetivos e qualitativos. Para o autor, a exclusão da consciência é um completo erro pois “a consciência é o fato central da existência especificamente humana, porque sem ela todos os outros aspectos da [...] nossa existência - linguagem, amor, humor e assim por diante - seriam impossíveis” (SEARLE, 2017, p. 21). A consciência, portanto, é fundamental para caracterizar as mentes e definitivamente precisa ser protagonista nas discussões sobre a mente.

Todavia, Searle nos alerta que para mudar esse cenário (de teorias que não incluem o estudo da consciência e dos aspectos subjetivos), é necessário desmistificar algumas visões tradicionais das quais o problema da consciência se torna mais latente. Esse é o caso, por exemplo, das teorias Dualistas e Materialistas que reforçam a oposição (ainda que por

diferentes maneiras) entre a natureza subjetiva da consciência e a busca por conhecimento objetivo da ciência. Para o autor, também é necessário que haja uma avaliação das concepções computacionais de mente porque elas apresentam uma exclusão metodológica total dos aspectos subjetivos e qualitativos da mente, sobretudo, a consciência.

Ademais, Searle considera que a “exclusão da consciência”, ou melhor, essa “negação do subjetivo” é consequência direta de ao menos dois importantes fatores: o primeiro deles é o rigor da objetividade metodológica imposta pelas ciências e o segundo é a supervalorização do pensamento materialista. Para o autor, o princípio científico de que a verdade depende exclusivamente do exame e da observação pública dos fenômenos (incluindo os mentais) fez com que qualquer teoria disposta a discutir os aspectos subjetivos e qualitativos fosse classificada como metafísica ou dualista. A bipartição do mundo (pela visão dualista) entre os fenômenos objetivos e os subjetivos é precisamente o que dificulta o estudo formal da consciência nos termos das ciências naturais. Porém, Searle (2006, p. 24) acredita que a solução para isso não é radicalizar as explicações materialistas, pois para o autor, isso nos leva ao engano de acreditar que as únicas escolhas de fato viáveis dão-se o materialismo ou alguma forma dualismo.

Obviamente, Searle considera a refutação do dualismo uma pauta de extrema importância, afinal muitos dos problemas tradicionais da filosofia da mente persistem até os dias de hoje porque os filósofos não conseguiram abandonar por completo a dicotomia dualista. No entanto, Searle acredita que é possível refutar o dualismo sem ter que excluir os aspectos subjetivos e qualitativos como fazem os materialistas. Para o autor, reconhecer esses aspectos não equivale a ser um dualista ou negar que a mente tenha de fato uma origem material.

Além disso, Searle diz que a radicalidade excessiva do materialismo não se aplica apenas à teoria dualista, mas a toda e qualquer forma de explicação do mental que não seja baseada na redução à realidade física, pois quando os materialistas declaram a identidade do mental com o físico parecem na realidade negar “que haja quaisquer propriedades fenomenológicas irreduzíveis, tais como a consciência e os qualia” (SEARLE, 2006, p. 44). Fato que torna algumas das explicações materialistas, ao menos para Searle, tão prejudiciais à noção de mente quanto as dualistas. Além disso, o radicalismo materialista favorece

também o surgimento de teorias e metodologias que excluem cada vez mais os aspectos subjetivos e qualitativos do estudo da mente.

Dentre essas teorias, Searle se mostra mais preocupado com aquelas que estão comprometidas com uma rígida noção reducionista, como é o caso das teorias materialistas de identidade, e sobretudo, das alegações advindas da metáfora computacional e do cognitivismo, uma concepção contemporânea de mente baseada na aproximação entre o computador e a mente.

Por certo, para o autor, o advento do computador e as ideias computacionalistas tornaram o problema da consciência ainda mais grave visto que as questões em torno da mente foram radicalmente dissociadas das noções qualitativas e intencionais para dar lugar às questões sintáticas, em especial, o estudo da cognição humana por meio de descrições computacionais. Searle diz que no Cognitivismo “os processos que explicam a cognição são inconscientes não somente de fato, mas em princípio. [...] não são a espécie de fenômenos que poderiam se tornar conscientes” (SEARLE, 2006, .282). E em razão disso, a posição de John Searle é completamente contrária. O autor ainda diz que a linguagem sintática dos programas de computador não é suficiente para caracterizar a mente humana, visto que a intencionalidade, uma de suas principais características, é essencialmente semântica.

De toda forma, sejam teorias advindas do materialismo, do dualismo ou do Computacionalismo, Searle considera o afastamento e a rejeição das noções subjetivas da mente um grave erro. Não por acaso, é sobre isso que ele discorre em sua famosa obra *The Rediscovery of the Mind* (A redescoberta da mente) de 1992, propondo uma análise das principais teorias do mental e conseqüentemente dos argumentos que favorecem ao longo da história o afastamento da consciência e dos aspectos subjetivos no meio científico.

A análise é elaborada sob o objetivo de apresentar uma alternativa que reverta esse quadro, isto é, que nos permita escapar das mesmas questões deixadas pelo problema mente-corpo. A alternativa apresentada pelo autor, sua teoria geral da mente, é o “Naturalismo biológico”. Nós não falaremos do Naturalismo Biológico em si, mas da discussão sobre a consciência presente nele. Para isso, decidimos apresentar não apenas os fundamentos da teoria da consciência do autor, mas também as objeções e críticas às teorias que (como mencionado) demonstram de forma mais radical o descaso com a consciência e os aspectos subjetivos. Dentre as discussões propostas por Searle em *A redescoberta da mente*, daremos

enfoque especial a três: a defesa da irreducibilidade das mentes; as objeções ao cognitivismo contidas em “A crítica da razão cognitiva” e por fim, a discussão sobre “o princípio da conexão”, uma tese em favor da centralidade da consciência.

A defesa da *irreducibilidade das mentes* implica numa crítica à noção reducionista de que a realidade subjetiva das mentes pode ser reduzida a descrições objetivas. Já *A crítica da razão cognitiva* apresenta severas objeções ao cognitivismo e sua alegação de que o cérebro é uma espécie de computador digital. Por sua vez, o *princípio da conexão* apresenta a tese searlina “[...] de que todos os estados intencionais inconscientes são em princípio acessíveis à consciência” (SEARLE, 2006, p.224). Esse princípio é usado pelo autor para estabelecer, dentre outras coisas, que não existem estados puramente inconscientes. O *princípio da conexão* não apenas constitui uma forma de demonstrar a importância da consciência para a compreensão da mente, mas em certa medida, “sintetiza” as razões pelas quais Searle é fortemente contrário ao cognitivismo e ao reducionismo, visto que essas metodologias acarretam a noção de que os fenômenos mentais são estados puramente inconscientes.

Por fim, para entendermos todas essas questões, organizamos este trabalho da seguinte maneira: o primeiro capítulo apresenta uma introdução geral das principais teorias e metodologias ligadas (ainda que em diferentes níveis) ao problema da consciência. No segundo capítulo, porém, apresentaremos a teoria geral da consciência de Searle, o que inclui, sua defesa da irreducibilidade das mentes e sua explicitação do princípio da conexão. Por último, o terceiro capítulo tematiza um compilado de críticas, em que são apresentadas tanto as objeções de Searle ao cognitivismo quanto às respostas de seus opositores.

## 2 O PROBLEMA DA SUBJETIVIDADE

O termo “subjetividade” é sem dúvidas o mais importante deste trabalho, afinal, Searle acredita que a subjetividade é o maior obstáculo a ser superado para que o problema da consciência, principalmente, no âmbito científico, seja resolvido. No entanto, para que isso ocorra é necessário o esclarecimento de alguns falsos juízos, principalmente, da noção geral de que não é possível discutir cientificamente a consciência porque sua natureza subjetiva se contrapõe à finalidade última da ciência, a saber, a explicação causal dos fenômenos sociais, físicos e mentais por meio exclusivo da objetividade ontológica.

Uma possível resolução do problema da consciência, portanto, requer antes de mais nada entender se existe ou não um conflito entre a realidade subjetiva da consciência e a objetividade das práticas científicas. Em razão disso, ao longo deste capítulo, veremos como Searle (que diz ter encontrado uma solução, cf. SEARLE, 2006) compreende e trabalha as questões em torno deste conflito. Para tal, apresentaremos um breve percurso sobre as metodologias e teorias que compõem (na visão de Searle) o conjunto de fatores que corroboram para o problema da consciência. Por conta disso, o conflito entre as noções de “subjetivo” e “objetivo” nesta discussão serão limitadas às teorias e metodologias ligadas ao trabalho de John Searle, em especial, aquelas que são diretamente citadas na obra *A redescoberta da mente*.

### 2.1 A objetividade científica

Na introdução do quarto capítulo da obra *A redescoberta da mente*, Searle questiona porque parece difícil conciliar o que sabemos sobre a consciência (por exemplo, que ela é uma experiência em primeira pessoa) com a nossa visão de mundo científica. Há de fato algo inconciliável entre a ciência e a consciência? Em resposta, o autor explica que nossa visão de mundo científica contemporânea tem dificuldades em aceitar essa conciliação porque herdou da modernidade a ideia de que a ciência deveria excluir tudo que fosse mental (subjetivo) e se ocupasse apenas da matéria<sup>1</sup>.

---

<sup>1</sup> Sobre essa questão o autor explica que: "nossa visão de mundo contemporânea começou a se desenvolver no século XVII, e seu desenvolvimento vem prosseguindo até o final do século XX. Historicamente, uma das chaves para esse desenvolvimento foi a exclusão da consciência do objeto da ciência por Descartes, Galileu e outros no século XVII.

Sem dúvidas, o emprego da objetividade nas questões metodológicas e epistemológicas foi uma das características mais marcantes da revolução científica ocorrida no início do século XVI e final do século XVIII. Neste período ocorreram grandes transformações no pensamento científico, principalmente, no que concerne à Cosmologia sob as descobertas do astrônomo Nicolau Copérnico publicadas e sua obra *Das revoluções das Esferas Celestes*. A descentralização da terra na ordem dos elementos celestiais não apenas expandiu a percepção dos homens sobre os horizontes do universo como também possibilitou o desenvolvimento de novas concepções matemáticas, geométricas, científicas e filosóficas. O resultado dessas transformações foi uma inevitável supervalorização da razão humana que concedeu à ciência um alto status de autoridade sobre aquilo que poderia ou não ser conhecido pelo homem.

Na obra *O que é a ciência afinal?* o filósofo Alan Chalmers explica que após a revolução científica a “[...] atribuição do termo *científico* [grifo nosso] a alguma afirmação, linha de raciocínio ou peça de pesquisa é feita de um modo que pretende implicar algum tipo de mérito ou um tipo especial de confiabilidade” (1993, p.17). Isso fez com que a ciência, ou melhor, o status de “cientificidade” se tornasse um requisito aspirado por quase todas as disciplinas e áreas do conhecimento humano. Além da supervalorização da razão, outra característica importante da revolução científica foi a formulação de métodos empíricos como critérios metodológicos considerados seguros, o que tornou as explicações científicas sobre os fenômenos e organismos exclusivamente empírica, já que a rejeição ou aceitação de uma hipótese dependia do resultado obtido em testes e experimentos<sup>2</sup>.

Também nascia a necessidade de garantir que esses métodos fossem de fato seguros, afinal, eles deveriam ser provas de validação para as novas hipóteses e teorias, isto é, deveriam ser formas seguras de determinar a validade de seu conhecimento. E para que isso fosse possível, o método deveria atender certas exigências, como, por exemplo, a neutralidade, a garantia de

---

Na concepção cartesiana, as ciências naturais verdadeiras excluía a “mente”, *res cogitans*, e ocupavam-se apenas da “matéria”, *res extensa*” (SEARLE, 2006, p. 126).

<sup>2</sup> O desenvolvimento de métodos científicos foi possível graças às novas tecnologias, como por exemplo, o microscópio e a luneta astronômica de Galileu Galilei. A tecnologia foi um dos fatores decisivos para que a nova formulação de ciência fosse cada vez mais objetiva, afinal, se entendia que a ciência deveria ser responsável por desenvolver os métodos e as técnicas necessárias para o aumento da qualidade e da produtividade do trabalho. Essa visão concedeu à tecnologia e a ciência uma relação intrínseca de influência mútua que transformou a ciência moderna em uma ciência da técnica. Ela se tornou técnica não apenas no sentido de produzir conhecimento prático, mas por ser capaz de intervir diretamente no progresso da sociedade. Fato este que mais tarde foi radicalizado na revolução industrial e no desenvolvimento de meios ainda mais eficientes de produção de trabalho.



reprodução e a intersubjetividade, pois um procedimento científico verdadeiramente eficiente não poderia ser realizado apenas uma única vez ou depender exclusivamente de uma única pessoa. Ao contrário, um método seguro teria de prescrever procedimentos reproduzíveis e observáveis por qualquer indivíduo capacitado para que assim não houvesse dúvidas quanto à sua eficácia. Dessa forma, além de oferecer um meio de verificação, o método deveria possuir uma garantia de validade. Mas como é possível garantir a validade de um método? O que pode ser considerado uma prova irrefutável?

Na filosofia, muitas respostas foram dadas a essa questão. Nomes como John Locke, David Hume e Immanuel Kant podem ser prontamente citados. No entanto, nos concentraremos na resposta oferecida pela tradição do positivismo lógico e seu famoso “princípio da verificação” que marcou toda uma geração de pensadores ao determinar com rigor aquilo que de fato era dever da ciência.

O positivismo é um movimento filosófico francês que remonta ao século XIX. Ele é considerado por muitos manuais como um movimento “fruto do iluminismo”. A razão disso é que o Positivismo conservou algumas das principais características do iluminismo, como por exemplo, o ideal de progresso através da ciência e da razão humana. A teoria positivista, assim como o iluminismo, influenciou toda uma geração de pensadores, tanto no meio científico quanto filosófico. O que por sua vez possibilitou que o positivismo fosse aplicado em diferentes teorias e vertentes. Dentre as vertentes surgidas, nos interessa comentar o positivismo lógico, dos filósofos, lógicos e matemáticos do círculo de Viena<sup>3</sup>.

A vertente do positivismo lógico conta com grandes nomes da filosofia como Hans Hahn, Otto Neurath e Rudolf Carnap. Ela foi fortemente influenciada pelo empirismo de David Hume e algumas das principais ideias da obra “Tractatus Logico-Philosophicus” de Wittgenstein. Como resultado, o positivismo lógico foi marcado pela supervalorização do empirismo na definição da ciência e na compreensão da linguagem, principalmente, das categorias de “significado” e “uso”. Algo expressado ao longo de todas as páginas do manifesto positivista intitulado “A concepção científica do mundo”. Uma passagem em particular nos apresenta com muita clareza o posicionamento adotado por eles em relação ao modo como a ciência deveria se comportar.

---

<sup>3</sup> O Círculo de Viena é o nome dado ao grupo de pensadores (que incluía matemáticos, filósofos e cientistas) que se reuniram na Universidade de Viena durante os anos de 1922 a 1936 para discutir questões sobre lógica, linguagem e ciência.

Caracterizamos a concepção científica do mundo essencialmente mediante duas determinações. Em primeiro lugar, ela é empirista e positivista: há apenas conhecimento empírico, baseado no imediatamente dado. Com isso se delimita o conteúdo da ciência legítima. Em segundo lugar, a concepção científica do mundo se caracteriza pela aplicação de um método determinado, o da análise lógica. (CARNAP, HAHN, NEURATH, 1986, p.12).

Ora, o que os autores estão dizendo é que a verdade científica não deveria ser construída pela mera formulação de hipóteses e ontologias especulativas. Ao contrário, qualquer enunciado científico teria de possuir uma correspondência com a natureza objetiva, pois somente através de uma verificação empírica é que se torna possível a confirmação de uma verdade.

De fato, segundo os positivistas lógicos, só são significativos os enunciados que podem ser verificados; é essencialmente isso que é declarado pelo famoso "princípio de verificação". Note-se que a exigência não é de verificação efetiva, mas sim a possibilidade de verificação. Além disso, pelo mesmo pensamento, para que um enunciado possa ser considerado científico, ele deve ser passível de verificação pelo método científico. Desse modo, o princípio de verificação também funciona como um princípio de demarcação entre aquilo que é científico e aquilo que não é, possibilitando assim, a clarificação dos conceitos apresentados pela ciência.

A classificação de cientificidade baseada na verificabilidade também tornou possível que os positivistas lógicos fizessem afirmações sobre o papel da filosofia, da linguagem e da ciência de modo que, os enunciados que não atendessem ao princípio verificacionista deveriam ser considerados sem sentido. O que, por consequência, fez com que os conceitos apresentados pela metafísica, ética e religião não fossem desconsiderados ciência. E para que o mesmo não ocorresse com os conceitos filosóficos, os positivistas lógicos concluíram que a filosofia deveria excluir toda e qualquer pauta metafísica, tornando a filosofia uma atividade comprometida exclusivamente com a análise lógica da linguagem.

Assim, se pode dizer que o positivismo lógico e o princípio verificacionista foram importantes fatores que reforçaram ainda mais o discurso construído no início da era moderna, a saber, o entendimento geral de que todos os fenômenos do mundo poderiam ser explicados e entendidos se estes forem submetidos à observação e à análise lógica. O apelo à objetividade é gerado, antes de mais nada, pelo ideal de que todos os fenômenos são formas de existência objetiva (isto é, possuem uma natureza que é objetiva) e, portanto, a verificação empírica deveria ser vista como uma prova final irrefutável.

Entretanto, segundo Searle, isso gera um problema para tudo aquilo que não pode ser submetido aos termos dessa concepção. A radicalização faz com que ela (a ciência) seja “[...] objetiva não somente no sentido de que se empenha por chegar a conclusões que sejam independentes de preferências e pontos de vista pessoais, mas, mais importante, envolve uma realidade que é objetiva. A ciência é objetiva porque a própria realidade é objetiva” (SEARLE, 2006, 20). Consequentemente, se cria um conflito com o estudo da consciência, afinal, ela é um fenômeno subjetivo e, portanto, segundo a concepção de ciência que vimos até aqui, a consciência não é um objeto viável para a ciência.

Em contraposição, Searle diz que essa é uma falsa concepção. Ele argumenta, em primeiro lugar, que “nem toda a realidade é objetiva; parte dela é subjetiva” (SEARLE, 2006, p. 32) e esse é o caso exclusivo da consciência. Em segundo lugar, ele diz que a objetividade científica não possui um real conflito com a subjetividade da consciência, pois ele acredita que a finalidade última da ciência não é explicar a existência objetiva das coisas, mas encontrar respostas objetivas sobre a existência delas. Porém, para sustentar esses argumentos, Searle acredita que existem certas questões que devem ser levadas em consideração, e a primeira delas é, sem dúvidas, a metodologia reducionista, uma consequência advinda do positivismo que dominou boa parte das teorias modernas de mente, em especial, as materialistas.

A metodologia reducionista se mostra um problema para a tese de Searle não apenas por reforçar o discurso sobre a objetividade científica, mas também por ser metodologicamente incompatível com a centralidade da consciência. Além disso, ela também contribui na perpetuação do discurso de que é possível determinar a natureza objetiva dos fenômenos, objetos e eventos (incluindo os mentais) se os submetemos ao processo de redução, que consiste (de forma geral) em reduzir continuamente as entidades e propriedades complexas a suas partes mais simples. Na verdade, as implicações desse princípio não ressoam apenas na ciência, mas também nas teorias filosóficas e psicológicas da mente. Mas, primeiramente, vejamos em que consiste a metodologia científica.

## **2.2 A metodologia reducionista**

O termo reducionismo corresponde à metodologia nascida no seio da ciência positivista e que consiste no uso de análises lógicas para demonstrar uma relação de equivalência entre um enunciado complexo e certa combinação de enunciados simples. Já em termos epistemológicos,

o reducionismo é o nome dado ao princípio científico de que o conhecimento verdadeiro dos fenômenos depende exclusivamente do conhecimento prévio de suas partes ou elementos constitutivos. Ou seja, o reducionismo prega a ideia de que algo (seja organismo, objeto ou fenômeno) é o produto de suas partes. Esse princípio metodológico pode ser identificado ainda no texto sobre a concepção científica do mundo.

[...] todo enunciado científico deve poder ser indicado por meio de uma redução a um enunciado sobre o dado, assim também o sentido de cada conceito, pertencente a qualquer ramo da ciência, deve poder ser indicado por meio de uma redução gradativa a outros conceitos, até aos conceitos de grau mínimo, que se relacionam ao próprio dado (CARNAP, HAHN, NEURATH, 1986, p.12).

A ideia empregada pelos positivistas a respeito do reducionismo é de que todo enunciado significativo é passível de redução e que toda forma de redução deve prosseguir até atingir os seus graus mínimos. Como um dos principais objetivos dos positivistas era trazer clareza aos conceitos filosóficos e científicos, o princípio reducionista surgiu como uma metodologia de esclarecimento da linguagem com a sugestão de analisar proposições complexas por meio de seus termos mais simples.

Ademais, embora o reducionismo positivista tenha nascido de discussões sobre a linguagem e problemas lógicos, sua metodologia foi capaz de transmitir uma ideia clara: é possível compreender algo através da análise de seus termos simples, ou como citado, através de seus “graus mínimos”. A aplicação desse princípio, por exemplo, em discussões gerais sobre a mente resulta na ideia de que os fenômenos e propriedades mentais de níveis mais elevados (os pensamentos, as crenças e as intenções) podem ser reduzidos a alguma forma de nível inferior (atividades neurais do cérebro). Mesmo sob o uso de diferentes objetos, como enunciados, fenômenos ou propriedades (sejam elas mentais ou não) a única exigência fundamental da metodologia reducionista é que o processo de redução permita algum tipo de verificação empírica ou ao menos a possibilidade dela.

De fato, a clara mensagem do princípio reducionista foi o que permitiu que ele fosse empregado para outras finalidades e por diferentes disciplinas. Algumas das vertentes surgidas foram: a redução ontológica, a redução ontológica de propriedades, a redução causal, a redução teórica e a lógica. Embora todas possuam relevância, no que concerne a este trabalho, apenas a redução ontológica e a redução causal são conteudisticamente importantes, pois a redução

ontológica é a vertente que Searle crítica mais severamente e a redução causal seria a vertente que mais se aproxima de sua concepção de mente (cf, SEARLE, 2006, p.166).

A redução Ontológica é possivelmente a mais comum dentre as três, pois é a vertente que em geral se encontra nas ciências naturais. Basicamente, a redução ontológica implica na ideia de que “todos os fenômenos complexos podem ser sempre reduzidos a explicações mais elementares em termos da natureza e do comportamento de seus componentes” (BENNET; HACKER, 2005, p. 389). Um exemplo que pode nos ajudar a compreender (e que é muito utilizado por reducionistas ontológicos) é o da “água”. Do ponto de vista da redução ontológica, a água nada mais é do que o  $H_2O$ . A fórmula “ $H_2O$ ”, por sua vez, significa a junção de dois átomos de hidrogênio com um átomo de oxigênio que são os elementos constitutivos encontrados na composição da água. Então, o que de fato pode caracterizar algo? Para a redução ontológica, são todos os elementos que constituem a composição dessa coisa (isto é, seu material elementar). Dessa forma, dizer que a água é  $H_2O$  é afirmar simplesmente que ela é ontologicamente constituída pela junção de dois átomos de hidrogênio com um átomo de oxigênio.

Por outro lado, a redução Causal propõe reduzir a causalidade dos fenômenos a eventos causais mais elementares. Esse tipo de redução é mais facilmente encontrado nas discussões de física quântica em que se tenta provar que “todos os processos causais no mundo cotidiano envolveriam eventos macroscópicos que poderiam ser considerados uma composição mereológica de eventos quânticos” (PESSOA, 2013, p.366). Para simplificar, deixaremos de lado a ideia de “eventos” e nos concentramos novamente no exemplo da água: em um nível macroscópico, a característica causal mais importante da água é a sua solvabilidade, a característica que permite que a água atue como um “solvente universal” capaz de diluir outros elementos ou componentes. Assim, quando certa substância solúvel é adicionada à água, pode-se dizer que sua dissolução é causada pela solvabilidade da água. Claro, em uma investigação reducionista que busque uma explicação formal desse processo procurará a causa da solvabilidade da água e essa causa será considerada como a causa primária da dissolução da substância. Em suma, a causa mais imediata é reduzida a uma causa mais básica. E qual é essa causa mais básica? No caso da água, é a polaridade da molécula  $H_2O$ , isto é, do fato da molécula atuar como ímã, de modo que, quando um composto iônico é adicionado à água, os ânions são atraídos pelos pólos positivos das moléculas de  $H_2O$ , e os cátions, pelos pólos negativos, quebrando as moléculas do soluto. Dessa forma, quando em contato com outros componentes, a

molécula H<sub>2</sub>O realiza a separação das substâncias ou novas ligações dependendo da carga elétrica de cada componente.

Em síntese, podemos dizer que a redução causal, ainda que diferente da redução ontológica, está de acordo quanto ao princípio geral de que é possível submeter a linguagem, os fenômenos, os organismos e os objetos a alguma forma de realidade objetiva, isto é, a alguma forma de realidade material. No entanto, no que concerne às discussões sobre a mente, em especial da consciência, a tarefa se torna muito mais complexa, visto que existe o problema da subjetividade: como reduzir os fenômenos mentais subjetivos aos processos físicos do cérebro?

Para Searle isso é impossível devido à natureza subjetiva dos fenômenos mentais. Na verdade, para o autor, “mesmo uma ciência perfeita do cérebro não levaria a uma redução ontológica da consciência da forma que nossa ciência atual pode reduzir calor, solidez, cor ou som” (SEARLE, 2006, p.168). Existe (como veremos na discussão da seção 3.1) uma barreira que a redução ontológica, por definição, não pode ultrapassar, e ela reside, justamente, na natureza subjetiva dos fenômenos mentais, em especial, da consciência. Por outro lado, veremos que Searle enxerga na redução causal um aparato metodológico mais adequado para a caracterização da mente.

O autor faz uso dessa metodologia sob a concepção de que a consciência é “uma característica emergente de determinados sistemas de neurônios da mesma forma que a solidez e a liquidez são características emergentes de sistemas de moléculas” (SEARLE, 2006, p. 162). Além disso, essa concepção é formulada por Searle com o objetivo de resolver (ou ao menos burlar) os problemas deixados pela redução ontológica da mente e pela noção geral do reducionismo que influenciou toda uma tradição de teorias tanto da filosofia como da psicologia. Antes, porém, devemos esclarecer melhor quais são as teorias às quais Searle se refere e quais são os problemas evidenciados por ele.

### **2.2.1 *O problema do dualismo e o reducionismo na Filosofia da Mente***

Enquanto a ciência se concentra em categorias como “objetivo” e “observável”, o estudo da mente (em especial da consciência) requer o reconhecimento de categorias como “subjetivo” e “qualitativo” termos que, no entanto, são considerados por alguns como "anti-científicos" por estarem fora do domínio da ciência (que é objetivo). Enquanto isso, na filosofia

a oposição entre as categorias “objetivo” e “subjetivo” foram fomentadas pelas concepções dualistas e seu entendimento de que as mentes pertencem a uma realidade imaterial.

De certo modo, para um sujeito do século XXI pode ser ridiculamente fácil se apegar a ciência e hastear uma bandeira “antidualista”, no entanto, apesar de parecer uma tarefa fácil, ela dificilmente seria, afinal, o dualismo, para além das questões formais da teoria filosófica, toca em um campo complexo e de difícil trato que é o senso comum.

Sob o ponto de vista do senso comum, acreditar na existência de um plano diferente da realidade (um espaço, lugar ou fenômeno) é algo simples e ordinário, principalmente, se estes forem expressados por termos religiosos como “pós vida”, “alma”, “paraíso” e “inferno”. É ainda mais comum para aqueles que possuem algum tipo de crença religiosa acreditar que a vida mental do homem tem lugar numa alma imortal que pertence a uma realidade diferente da objetiva. Dessa forma, o simples fato de aceitar que existem duas formas de realidade distintas, é em certa medida, concordar com o princípio dualista de que a realidade pode ser entendida por mais de uma maneira. Há, no entanto, diferenças entre o dualismo do senso comum e o dualismo que vemos na filosofia.

Por mais que o dualismo filosófico admita que o mundo é constituído por duas realidades diferentes, essas realidades não são necessariamente criações de Deus ou deuses (embora em alguns casos essa possa ser a sugestão). Na Grécia antiga, Platão defendia a crença na existência de um mundo ideal onde habitavam as ideias perfeitas e eternas e de um mundo sensível onde a realidade material estaria em constante movimento (transformação). O homem não era diferente do mundo, seu corpo também estava em transformação, mas sua alma não. A mente, ao contrário do corpo, era parte do mundo perfeito, eterno e imaterial.

Já na modernidade, essa bipartição do mundo ganhou força através do *Dualismo de Substâncias* de René Descartes. Assim como Platão, Descartes também acreditava que o mundo era regido por duas ordens diferentes, porém, Descartes deu um passo a mais quando designou a mente do homem como “consciência” (ou autoconsciência) e a separou do cérebro (corpo), transformando essa relação em um problema filosófico que permanece sem resposta até os dias de hoje. Vejamos agora alguns pontos resumidos dessa teoria.

A hipótese de Descartes implica na existência de duas substâncias constituintes: a matéria comum (res-extensa) responsável por preencher o espaço, os objetos e as coisas do mundo físico. E a substância imaterial (res-cogitans) que se ocupa da atividade do pensar e que

constitui a alma<sup>4</sup>. Embora ambas as substâncias sejam importantes para a constituição do mundo e sejam capazes de interagir, o autor acredita que a *res-cogitans* pode existir independentemente da matéria física. Há dois argumentos interessantes a serem mencionados, são eles: o da dúvida e o da divisibilidade.

Em primeiro lugar, para Descartes o homem poderia duvidar de qualquer coisa no mundo, exceto de sua própria existência como um ser pensante. Mais precisamente, o autor argumenta que o pensamento é a única coisa no mundo da qual o homem não poderia duvidar. Ora, alguém que se questiona “eu penso?”, já não está pensando? Desse modo, duvidar do próprio pensamento é impossível. Por outro lado, parece possível supor que qualquer sujeito pensante pode perfeitamente duvidar da existência de seu próprio corpo. A grosso modo, a conclusão apresentada pelo autor é de que a existência da coisa pensante é indubitável e a existência do corpo é dubitável, o que parece implicar que são duas coisas diferentes.

Outro argumento utilizado por Descartes é o da “divisibilidade”. A hipótese desse argumento pode ser expressa pelo seguinte silogismo: “o corpo é uma extensão material e pode ser divisível; a mente é uma substância imaterial e por isso é indivisível, logo o corpo e a mente não podem ser iguais”. Ora, se os objetos materiais são extensos (ocupam espaço), seus corpos estão suscetíveis à separação e a divisão em partes. Para Descartes, é possível dividir qualquer corpo (coisa extensa) em outras partes, isto é, corpos menores. No entanto, o mesmo não pode ser feito com a mente. Sua existência imaterial (subjéctiva/espiritual) não permite que ela seja dividida da mesma forma como ocorrem com os corpos e os objetos. Sendo assim, a conclusão do autor é de que corpos e mentes não podem ser a mesma coisa.

Claro, alguns podem argumentar que o que torna o argumento da divisibilidade válido para os leitores é a pré-aceitação de que a mente seja de fato imaterial. Porém, acreditamos que um exercício que pode nos ajudar a entender o argumento de Descartes é questionar: é possível separar ou dividir os fenômenos mentais, como por exemplo, a intencionalidade e a memória, do restante da mente? A resposta mais intuitiva parece ser não.

De toda maneira, ainda que haja divergências quanto aos argumentos do autor, não se pode negar que a divisão do mundo em duas realidades bem como a distinção ontológica entre

---

<sup>4</sup> Assim como em algumas religiões, no dualismo cartesiano, a palavra “alma” significa “mente” ou “consciência”, ela é o elemento pensante do homem. A *res-cogitans*, é, portanto, a substância imaterial que compõe a natureza da mente humana.



aquilo que é mental e físico foram um dos grandes responsáveis pela popularização do problema “mente-corpo” na filosofia da era moderna. Na verdade, a complexidade das implicações decorrentes dessa questão, bem como o fracasso de Descartes em explicar a interação mente-corpo, fez com que muitos filósofos abandonassem o dualismo de substância pelo monismo materialista.

O monismo do tipo materialista entende que tanto a mente, o corpo e os objetos do mundo são constituídos por apenas uma única substância que é a matéria física. Assim, para o monismo materialista, “[...] nossa vida mental nada mais seria do que uma grande variação dos estados químicos e físicos de nosso cérebro. Nossas angústias, desejos e intenções seriam apenas um produto do cérebro e supor que tenham existência autônoma não passaria de uma ilusão” (TEIXEIRA, 2011, p.65). Dessa forma, não há razões para supor, como sugeriu Descartes, a existência de uma *res-cogitans* ou qualquer outra substância que esteja além do plano material.

Além disso, apesar dessas teorias serem baseadas em metodologias reducionistas para fundamentar suas hipóteses, no entanto, elas apresentam diferenças significativas. Para exemplificar, mencionamos as duas principais teorias do monismo materialista. São elas: o materialismo reducionista e o materialismo eliminacionista.

Como explica Teixeira, o materialismo reducionista (também conhecido como “teoria da identidade”) defende a hipótese de que “[...] os estados mentais são idênticos a estados cerebrais ou estados do sistema nervoso”. Em outras palavras, essa teoria poderia ser resumida através da seguinte equação: estados mentais = estados cerebrais” (2011, p.70). Isto é, os fenômenos mentais correspondem, de alguma maneira, a um processo cerebral (processo físico) que é numericamente idêntico ao fenômeno mental<sup>5</sup>.

Os defensores dessa teoria acreditam que um dia a ciência será capaz de demonstrar como efetivamente essa relação é estabelecida. Não por acaso, foram justamente os avanços científicos no campo de estudo da mente o ponto determinante para haver entusiasmo com a crença em uma relação de identidade. Principalmente, após a criação dos métodos para capturar a atividade do cérebro em imagem (neuroimagem). Na verdade, desde a descoberta da capacidade do cérebro em gerar atividade elétrica feita pelo alemão Hans Berger, abriu-se um intenso debate a respeito dos processos que ocorrem na estrutura do cérebro. Contudo, o desafio

---

<sup>5</sup> Segundo Leclerc, a noção de identidade do materialismo prever que “[...] um tipo de mental (type) ou propriedade mental (sentir dor, ou, bem-estar etc.) é idêntico a um tipo neural ou propriedade do cérebro (ter tal grupo de neurônios disparando, produzir tal neurotransmissor etc.)” (LECLERC, 2018, p.75).

ainda hoje é tornar possível o registro total dessas atividades.

Com isso, o que resta para a teoria da identidade é esperar que a neurociência possa demonstrar uma correspondência total que seja bem-sucedida. Isso parece ser algo compartilhado pelas teorias materialistas, visto que o materialismo do tipo eliminacionista também acredita que a neurociência será capaz de explicar como ocorrem os fenômenos e os processos cerebrais, oferecendo respostas verdadeiras sobre a mente. E mais importante ainda, isso será feito com base em novos termos científicos, e não nos termos mentalistas usados pela tradição dualista.

Para o materialismo eliminacionista, os termos “mentalistas” estabelecidos pelas teorias dualistas e pelo senso comum (ou psicologia popular) são os grandes responsáveis por “mistificar” a concepção de mente. Para essa vertente, o uso de termos, como por exemplo “alma” para caracterizar a ideia de mente apenas reforça a concepção de que o corpo e a mente possuem naturezas distintas.

Paul Churchland, um dos fundadores dessa teoria, explica que o erro dualista foi justamente usar a linguagem ordinária (do senso comum) para conferir à mente um estatuto ontológico que ela não possui. Para o autor, o “[...] senso comum é uma concepção falsa e radicalmente enganosa das causas do comportamento humano e da natureza da atividade cognitiva” (CHURCHLAND, 2004, p. 55). Em razão disso, a proposta Eliminativista consiste em eliminar os termos mentalistas, tanto os que se originam do senso comum quanto os estabelecidos pela tradição dualista, pois insistir no vocabulário mentalista é insistir nos mesmos erros cometidos pelos dualistas. Até que a neurociência seja capaz de apresentar respostas satisfatórias sobre o real “arcabouço mental”, o materialismo eliminacionista insiste que tanto a ciência quanto a filosofia devem se afastar das ideias e conceitos utilizados pelo senso comum que erroneamente são empregadas por teorias especulativas. Isso evitará conflitos desnecessários para a filosofia da mente, como por exemplo, o famoso problema mente-corpo.

Por outro lado, Searle acredita que assim como o materialismo reducionista, o materialismo eliminativista não resolve a dicotomia dualista e se mostra ainda mais radical. A impressão do autor é que “no caso do materialismo eliminativo, o que se considera como a justificativa da eliminação desses estados mentais em favor de uma neurobiologia perfeita é exatamente o pretense fracasso de qualquer correspondência semelhante” (SEARLE, 2006, p.74). Na verdade, a forma como a teoria materialista torna os aspectos subjetivos

metodologicamente irrelevantes apenas reforça o discurso de que a consciência não é adequada para o estudo científico. No fim das contas, Searle enxerga o materialismo e o dualismo como duas faces da mesma moeda. Eles perpetuam o mesmo problema: a incompatibilidade entre os fenômenos mentais e a realidade física.

Como veremos a seguir, o behaviorismo tampouco apresenta soluções melhores. Na visão de Searle, ele é apenas mais uma teoria da emaranhada rede de teorias aprisionadas na dicotomia físico-mental.

### ***2.2.2 O reducionismo na Psicologia e o aspecto comportamental***

Também podemos encontrar a influência da metodologia reducionista na psicologia, como por exemplo, nas ideias por trás da famosa teoria Behaviorista criada por John Broadus Watson no início do século XX. A teoria behaviorista surgiu com o principal objetivo de tornar a psicologia uma ciência rigorosa capaz de atender às exigências científicas e positivistas de sua época. Para realizar esse propósito, Watson deduziu que a psicologia deveria seguir um método científico que permitisse a observação do seu objeto de estudo da mesma forma como era feito nas ciências naturais. Porém, havia um problema: como a mente poderia ser submetida aos métodos tradicionais da ciência e do positivismo, como por exemplo, o verificacionismo?

A solução apresentada pelo autor foi de abandonar a mente (que não se caracteriza como um objeto observável) e se concentrar em suas relações causais, mais especificamente, aquelas manifestadas através do corpo, isto é, o comportamento observável. A hipótese de Watson era de que o comportamento (a partir de alguns estímulos em certas situações e em certos ambientes) poderia indicar determinadas atividades mentais, como por exemplo dor, felicidade e prazer. Ou seja, Watson propôs que seria possível estabelecer uma redução causal entre os aspectos mentais e os comportamentais. De modo que “A mudança de concepção metodológica apresentada pelo behaviorismo foi muito importante para a psicologia porque permitiu o abandono da introspecção como a principal ferramenta de trabalho do psicólogo, algo visto como uma vitória científica, visto que o autorrelato já não era considerado uma fonte confiável de investigação por se tratar justamente de uma descrição subjetiva que não fornecia maneiras seguras de verificação ou validação.

Ademais, segundo Churchland (cf. CHURCHLAND, 2004), existem três fatores que foram essenciais para a popularização da teoria behaviorista. Em primeiro lugar, havia a forte

necessidade de apresentar uma resposta satisfatória ao dualismo. Um segundo fator é que a teoria apresentava forte influência metodológica do positivismo lógico e do conceito de verificabilidade, o que tornou essa teoria preocupada em apresentar, sobretudo, um objeto de estudo que fosse empiricamente observável. Em terceiro lugar, está a defesa da ideia de que as questões deixadas pelo dualismo e pelo problema mente-corpo são na realidade problemas linguísticos que deveriam ser resolvidos por uma análise mais rigorosa dos conceitos e dos termos utilizados para se referenciar a mente.

Os fatores citados por Churchland são atribuídos à teoria behaviorista como um todo (isto é, de maneira geral), porém, acreditamos que elas se tornam mais visíveis quando analisamos suas vertentes. Por exemplo, no primeiro caso, que envolve uma resposta ao dualismo, podemos destacar as contribuições da abordagem do behaviorismo ontológico. Segundo Leclerc (cf. LECLERC, 2018, p.65), o behaviorismo ontológico compreende que tanto os atos, quanto os eventos e estados mentais de primeira pessoa são ilusórios. Para essa abordagem, não existe uma mente imaterial como afirmam os dualistas, na verdade, sequer existe uma. Para um behaviorista ontológico, os estados mentais são apenas reações comportamentais decorrentes da interação dos seres com o meio ambiente.

Por outro lado, o behaviorismo metodológico (o tipo apresentado por Watson) não nega a existência dos estados mentais, pelo contrário, ele afirma que eles existem. No entanto, a questão é encontrar uma solução para o problema metodológico em torno da mente. Ora, se a mente não é um objeto observável, é necessário que a psicologia encontre uma maneira de tornar possível a análise dos estados mentais, uma vez que a verificabilidade empírica é uma exigência científica.

Como já mencionamos, Watson propõe o estudo comportamental, contudo, essa troca de objetos, ou melhor, o uso de uma metodologia baseada na redução causal implica (consequentemente) na concordância de que os aspectos mentais subjetivos são irrelevantes para o estudo formal da mente, visto que para essa abordagem seria possível determinar certos estados mentais apenas pela análise das atividades causais expressas pelo comportamento. Logo, embora o behaviorismo metodológico não negue a existência dos estados mentais, ele tampouco os considera metodologicamente importantes<sup>6</sup>.

---

<sup>6</sup> O behaviorismo metodológico de Watson também foi fundamental para o surgimento do behaviorismo radical defendida pelo psicólogo Burrhus Frederic Skinner que por sua vez, apresentou importantes contribuições à psicologia ao afirmar que nossos comportamentos psicológicos são uma espécie de resposta produzida pelos estímulos que

Já o terceiro e último ponto mencionado por Churchland, se torna mais aparente nas questões do Behaviorismo lógico (ou filosófico) originalmente criada pelo britânico Gilbert Ryle (1900–1976) e difundida na comunidade filosófica. Em todo caso, ainda que tenha berço na filosofia, essa vertente conserva algumas das principais características do behaviorismo tradicional, como por exemplo, sua oposição ao dualismo. O principal propósito de Ryle com o behaviorismo lógico é propor uma análise gramatical do vocabulário usado para se referir sobre a mente e os fenômenos mentais. O objetivo do autor era demonstrar que os termos mentalistas (difundidos sobretudo pelo dualismo e o materialismo) eram na verdade expressões enganadoras que nos induziam inevitavelmente a pensar a mente sob o ponto de vista ontológico.

Na obra “O conceito de mente” publicada em 1949, Ryle crítica abertamente as ideias de Descartes afirmando que a linguagem e os termos dualistas utilizados para se referir à mente são, na realidade, erros de categoria. Esse problema é evidenciado na obra através do seguinte exemplo:

A um estrangeiro que visite pela primeira vez Oxford ou Cambridge, mostramos um certo número de faculdades, bibliotecas, campos de jogos, museus, departamentos científicos e secretarias administrativas. Então ele pergunta: “Mas onde está a Universidade? Vi onde vivem os membros das faculdades, onde trabalham os funcionários, onde os cientistas fazem experimentos [...]. Mas ainda não vi a Universidade em que residem e trabalham os membros da vossa Universidade (RYLE, 1970, p.18).

A partir dessa situação, Ryle explica que o erro cometido pelo estrangeiro é o mesmo tipo de erro cometido pelo dualismo. Assim como o visitante acredita que o conceito “universidade” significa algo além (separado) do conjunto de edifícios, locais e departamentos mostrados, o dualista acredita que “mente” se refere a algo além do mundo físico. No caso do visitante, Ryle diz que é necessário explicar o óbvio: a universidade não é nada além do que ele já viu. Ela é, precisamente, a maneira como todos aqueles setores funcionam e são organizados.

O autor prossegue a discussão afirmando ainda que o erro do visitante é uma consequência direta da “ousadia” filosófica de supor que há uma forma subjacente de existência que está além do comportamento dos seres vivos e do meio físico observável. Essa forma subjacente mencionada por Ryle é justamente a substância imaterial de Descartes. Pois ele

---

recebemos desde o nosso nascimento. Não por acaso, o autor inicia o primeiro capítulo de sua obra “O comportamento verbal” com a frase: “os homens agem sobre o mundo, modificam-no e, por sua vez, são modificados pelas consequências de suas ações” (SKINNER, 1978, p.15).

acredita que postular conceitos relativos a substâncias ou entidades imateriais é um erro, porque os conceitos são, em primeiro lugar, formulações que partem da reunião de propriedades ou características observáveis, o que não é o caso da res-cogitans.

No livro “Mente, cérebro e cognição”, Teixeira apresenta o conceito de “casamento” para exemplificar a teoria de Ryle. O autor diz (em uma versão resumida) que o conceito de casamento designa um tipo de relação que não é palpável e nem observável, mas que é extraída da observação de suas consequências causais na vida das pessoas envolvidas. Não se trata de uma entidade casamento, mas de um conceito que foi construído a partir da observação de suas propriedades causais. Logo, Ryle conclui que nem a linguagem nem o comportamento são suficientes para conceber a mente como um tipo de substância ou fantasma na máquina<sup>7</sup>. Ambos, na realidade, podem apenas indicar ou referenciar suas relações causais.

Apesar do behaviorismo representar mudanças importantes para a Psicologia, a teoria recebeu severas críticas. Searle, por exemplo, diz que o problema geral do behaviorismo “repousa no fato de que ele nega a existência de quaisquer estados mentais internos além do comportamento exterior” (SEARLE, 2006, p.54). Por esse motivo, Searle rejeita tanto o behaviorismo quanto o dualismo e o materialismo. Apesar da imaterialidade das mentes ser um completo erro para Searle (e por isso, para ele, é impossível defender o dualismo), isso não o faz simpatizar com o materialismo ou o behaviorismo, afinal, essas teorias negam verdades que são óbvias, como por exemplo, a realidade das experiências subjetivas.

Diante da perspectiva apresentada, conclui-se que há uma clara necessidade de encontrar uma explicação materialista que de fato possa explicar com clareza as relações entre os estados mentais e o cérebro. Tal explicação parecia começar a se insinuar no início da década de 50, tão logo os trabalhos de alguns pensadores interessados em computação proporcionaram uma nova concepção de mente baseada em analogias entre computadores, mentes e cérebros. Com o Computacionalismo prevaleceu a ideia de que a mente humana é um tipo de sistema de processamento de informações e que funciona de maneira similar à computação realizada pelos programas computacionais. Outra importante mudança é que as teorias psicológicas baseadas no Computacionalismo passaram a tratar a relação “mente- corpo” por meio da analogia “software-hardware”.

---

<sup>7</sup> Termo utilizado por Ryle para exemplificar a relação dualista de mente-corpo. Segundo o autor, a mente imaterial dualista habita o corpo físico como um fantasma preso em uma máquina.

## 2.3 A Metáfora Computacional

Ainda que a ideia de computação já fosse tema de discussões entre matemáticos e engenheiros, foi somente a partir de 1946 com a criação do ENIAC<sup>8</sup> que houve um grande interesse a respeito dos computadores, principalmente, pelas possibilidades de uso científico dessas máquinas inteligentes. No estudo da mente, as aproximações entre os computadores e a cognição humana geraram uma nova concepção de mente, dando início a chamada "metáfora computacional" que marcou os anos de 1940 a 1970.

A popularização da Metáfora computacional ocorreu por causa do reconhecimento de pelo menos dois paralelos principais entre a mente e o computador: primeiro, havia a tese geral de que as mentes eram semelhantes aos computadores na medida em que as mentes são uma espécie de sistema cuja principal função é o processamento de informação. Tanto as mentes quanto os computadores processam informações de maneira simbólica, isto é, em um nível representacional. Em segundo lugar, também havia a constatação de que os computadores, assim como as mentes, possuíam dois níveis de interface, o software e o hardware. E assim, as comparações entre os níveis do computador e os da mente proporcionaram um paralelo entre as mentes e os softwares, os "níveis representacionais", e entre os cérebros e os hardwares, os "níveis físicos".

Ademais, além das alegorias citadas, existiriam outros fatores determinantes para o sucesso da metáfora computacional, e dentre eles os mais importantes foram: o esforço dos matemáticos em construir sistemas simbólicos baseados em regras lógicas de representação do conhecimento, o surgimento da inteligência artificial, as críticas à teoria behaviorista somada com a necessidade de superar os problemas do critério comportamental, a superação dos modelos simbólicos em favor dos modelos neurais e pôr fim a revolução cognitiva que trazia a biologia como a disciplina central para entender a relação entre a estrutura do cérebro e a mente.

### 2.3.1 *Mentes simbólicas e a Inteligência Artificial*

Searle, que por muito tempo se considerou um cientista cognitivo, diz que as ideias iniciais que instigaram a metáfora computacional foram melhor expressas nos trabalhos dos

---

<sup>8</sup> O ENIAC é a sigla para "Computador e Integrador Numérico Eletrônico". Trata-se do primeiro computador digital do mundo.

matemáticos Alonzo Church e Alan Turing. Claro que Church e Turing não foram os únicos responsáveis pela consolidação da metáfora; nomes como John Von Neumann, Karl Lashley, Warren McCulloch e dentre outros são de extrema importância para essa temática. No entanto, tendo em vista que o trabalho de Turing é mais expressivo na filosofia da mente e porque o autor é muito citado por Searle, daremos uma atenção especial aos resultados alcançados por ele, principalmente, ao seu modelo de máquina computacional e ao seu famoso teste, o “jogo da imitação”.

Quando ainda era um estudante em Cambridge, Turing se interessou pelo problema da decisão levantado pelo alemão David Hilbert que, em linhas gerais, consistia em saber se era possível ou não criar um método preciso para decidir a validade de proposições matemáticas. O método deveria fornecer uma resposta “sim” ou “não” para problemas como: o número  $y$  é divisível por  $x$ ? Ou  $x$  é um número primo? A principal preocupação de Hilbert era então saber se:

[...] existe um método definido que, pelo menos em princípio, pode ser aplicado a uma dada proposição para decidir se ela pode ser provada. Num cálculo restrito, tal método pode de fato existir: por exemplo, a técnica da tabela de verdade para decidir se uma fórmula da lógica proposicional elementar é uma tautologia. Poderia existir um procedimento de decisão como este para proposições matemáticas? (HODGES, 2001, p. 14).

Em seu artigo “On computable numbers with an application to the entscheidungsproblem” de 1936, Turing apresentou um modelo de máquina que seria capaz de computar qualquer problema passível de ser descrito sob uma regra lógica interna. Mais tarde a máquina foi chamada de “a máquina de Turing” e a regra de “algoritmo”. Mas o que exatamente é um algoritmo e uma máquina de Turing? Quais foram os resultados alcançados por ele?

O algoritmo formulado por Turing é uma regra lógica capaz de realizar uma determinada tarefa através de uma sequência finita de comandos simples e sequenciais. Ou seja, a partir de um conjunto de instruções simples e precisas é possível a realização de tarefas mais complexas. Já a máquina de Turing se trata de um modelo teórico construído para rodar o algoritmo. Sua estrutura possui uma “fita” que é dividida em seções (chamadas de quadrados) e um scanner que é encarregado de ler os símbolos impressos nos quadrados da fita. Dessa forma, o scanner lê os símbolos na fita e cada um deles determina um comando específico que a fita e



o scanner devem realizar. Por exemplo, o símbolo “0” ordena que o scanner se mova dois quadrados para a esquerda, já o símbolo “1” ordena que ele avance um quadrado para a direita. Obviamente, os comandos citados são apenas exemplos simples do que poderia ser feito. O importante, porém, é entender que “na medida em que se pudesse expressar claramente os passos necessários para a execução de uma tarefa, ela poderia ser programada e executada pela máquina de Turing, que simplesmente leria a fita (independentemente de seu comprimento) e seguiria as instruções” (GARDNER, 2003, p.32).

É importante esclarecer que a máquina de Turing é basicamente o modelo teórico do computador digital. De fato, seguindo as definições de Turing, "um computador digital pode ser usualmente visto como constituído de três partes: a) memória b) Unidade executiva c) Controle" (cf, TURING, 1973, p. 54). A memória tem uma função semelhante ao papel que usamos para fazer cálculos. Nele podemos ver as fórmulas e os passos dados para a resolução do problema.

Quando o computador humano (a mente de um indivíduo) realiza um cálculo de cabeça, a memória desempenha o mesmo papel. Paralelamente, é isso que vemos no modelo da máquina de Turing, em que a fita desempenha o mesmo papel, isto é, armazenar as regras (Turing chama isso de livro de regras). A unidade executiva, por sua vez, é responsável por realizar as operações determinadas pelo livro de regras e o controle assegura que essas regras serão obedecidas seguindo a ordem correta.

A questão é que a mente humana também parece ter características de um sistema que realiza tarefas por meio de operações executadas sequencialmente, isto é, características de um computador digital. Assim, apesar do foco de Turing ser inicialmente a resolução de problemas matemáticos, sua noção de uma máquina computacional proporcionou importantes consequências para o estudo da cognição. De modo especialmente marcante, ela contribuiu para reforçar a ideia de que as mentes são de fato sistemas de processamento que operam em um nível simbólico e são implementados em sistemas físicos, isto é, em cérebros. A ideia de IA reside justamente no pressuposto de que tarefas cognitivas podem ser realizadas por um sistema artificial, desde que o programa certo seja implementado.

A aceitação acadêmica do trabalho inicial de Turing o levou, posteriormente, a se questionar sobre a capacidade inteligente dos computadores em seu artigo “Computing Machinery and Intelligence” publicado em 1950. Embora o termo Inteligência Artificial tenha sido oficialmente criado e utilizado por John McCarthy em uma conferência em 1957, é nesse artigo

de Turing que se encontram as bases da IA. Nele, pela primeira vez, um pensador “ousou perguntar se uma máquina pode[ria] pensar” (HODGES, 2001, p. 9). Na realidade, Turing sabia das dificuldades e implicações envoltas no uso do termo “pensar” e foi por essa razão que propôs o "jogo da imitação", um teste hipotético capaz de medir ou ao menos reconhecer as capacidades linguísticas dos computadores. O teste, que posteriormente ficou conhecido como o teste de Turing, era formado por perguntas e respostas que deveriam medir a habilidade do computador em simular a mente humana, ou seja, o teste visava analisar se o raciocínio usado pelo programa computacional podia ser comparado ao raciocínio humano.

O jogo, se descrito de maneira geral, ocorreria da seguinte forma: em um local designado, algumas questões seriam lançadas para dois indivíduos separados em cabines. Um deles é um ser humano e o outro é um computador. Ambos não teriam suas identidades reveladas e deveriam responder às questões através de “sim ou não”. Um terceiro indivíduo (um especialista) seria encarregado de fazer as perguntas e ao final identificar os indivíduos. Caso não fosse possível distingui-los, o computador passaria no teste, imitando com perfeição a inteligência humana.

Considerando a motivação do teste de Turing, podemos concluir que, embora o jogo da imitação tenha um caráter linguístico, a inteligência que ele pretende avaliar é uma inteligência de propósito geral, ou seja, uma inteligência de alta complexidade, uma que pode realizar as mais variadas tarefas cognitivas. Não obstante, existem IA 's mais simples, como por exemplo, o reconhecimento facial dos smartphones. Mesmo aplicações mais complexas, como é o caso dos softwares chatbot que são capazes de conversar e resolver problemas dos usuários em, por exemplo, um site de compras (os chamados atendentes virtuais), estão abaixo do nível de inteligência visado pelo teste de Turing. Um caso mais controverso, porém, é o recente caso do Chat GPT, provavelmente o chatbot mais famoso da atualidade. Diferente de outros atendentes virtuais, seu processamento de linguagem natural o permite gerar respostas em uma linguagem que simula de forma muito semelhante uma conversa humana. É possível então supor que, dadas as condições do jogo da imitação, o Chat GPT3, ou pelo menos o Chat GPT4, passaria no teste de Turing. Por outro lado, pode-se argumentar que o sucesso do programa no teste dependerá do nível de rigor do entrevistador. Um entrevistador suficientemente rigoroso e competente não seria enganado por esse tipo de IA.

Naturalmente, o trabalho de Turing também recebeu críticas, principalmente, a confiabilidade de seu teste (o quarto chinês, que é um argumento de Searle a ser examinado mais

à frente, é um exemplo disso), no entanto, as críticas não podem negar o fato de que suas ideias foram essenciais para o sucesso da metáfora computacional e para a criação de novas disciplinas, como por exemplo, a Inteligência Artificial (IA). A própria Filosofia da Mente acabou sendo afetada por essas ideias e o resultado disso foi o surgimento de teorias comprometidas em explicar o funcionamento da mente a partir da lógica computacional. Dentre elas, a mais famosa é a teoria Funcionalista cujos principais representantes são Daniel Dennett, David Lewis, Hilary Putnam e Jerry Fodor. O princípio norteador adotado por esses autores é de que os estados mentais não são estados cerebrais no sentido do materialismo reducionista, mas sim estados funcionais que são realizados por estados cerebrais.

### 2.3.2 O Funcionalismo

O funcionalismo surgiu em meados de 1960 como uma tentativa de oferecer uma alternativa ao fracasso behaviorista em explicar os estados mentais com base no aspecto comportamental e à implausibilidade da tese da identidade entre os estados mentais e os estados cerebrais postulada pelas teorias da identidade.

Hilary Putnam, fundador da teoria, caracterizava a relação entre a mente e o cérebro através da analogia computacionalista de “software-hardware”. O autor era simpático às ideias de Turing porque acreditava que a lógica por trás de sua máquina era fundamental para entendermos o funcionamento dos estados mentais. Além disso, Putnam também era um entusiasta da IA, pois considerava que o sucesso desse programa poderia corroborar a ideia de que a análise mais adequada do problema mente-corpo é a análise funcional.

Outra figura muito importante para a teoria funcionalista é o discípulo de Putnam, Jerry Fodor. Seu livro "Language of Thought", onde defende que o pensamento funciona como uma linguagem inata, é considerado um dos grandes clássicos da filosofia da mente e das ciências cognitivas. De fato, a hipótese da linguagem do pensamento é considerada (junto com o trabalho de Turing) um dos grandes alicerces do pensamento computacionalista.

Como um funcionalista, Fodor entendia que:

O Funcionalismo é a filosofia da mente baseada na distinção que a ciência da computação traça entre o hardware do sistema, ou composição física, e seu software, ou programa. A psicologia de um sistema como um ser humano, uma máquina ou um espírito desencarnado não dependeria do material de que é feito

o sistema (neurônios, diodos ou energia espiritual), mas do modo como esse material é organizado. (FODOR, 2011, p.10)

Essa passagem apresenta justamente uma das teses mais notórias do funcionalismo, a *tese da múltipla realizabilidade*, segundo a qual, da mesma forma que um programa, um mesmo estado mental pode ser realizado em diferentes tipos de *hardware*. Essa tese é fundamental para a refutação do materialismo reducionista, pois ela estabelece que a identidade entre estados mentais e cerebrais é impossível, visto que o mesmo estado mental pode ser implementado nos mais diferentes substratos, inclusive em um espírito desencarnado, o que aponta para a neutralidade do funcionalismo no debate entre fisicalismo e dualismo. Não por acaso, essa é a principal razão para a mente dos funcionalistas ser descrita como “similar a uma máquina de Turing”. Em suma, os pressupostos do Computacionalismo clássico de Turing e a teoria funcionalista de Putnam e Fodor forneceram as bases filosóficas da revolução cognitiva, isto é, a revolução que deu origem às ciências cognitivas.

### 2.3.3 A Ciência Cognitiva e o surgimento do Cognitivismo

As novas teorias da mente e da linguagem, alinhadas com as novas propostas metodológicas da psicologia (após o fracasso dos behavioristas) proporcionaram o surgimento da “Ciência Cognitiva” e do “Cognitivismo”. De fato, desde a introdução deste trabalho, fizemos uso vez ou outra desses termos; neste ponto, porém, se faz necessário esclarecer quais são as diferenças entre os seus significados e usos.

O termo “ciência cognitiva” se refere a uma disciplina nascida no século XX que se ocupa em explorar os aspectos da cognição humana, como por exemplo, o aprendizado, a percepção, a memória e a linguagem. Uma forma de descrevê-la é oferecida por Gardner da seguinte maneira: “a ciência cognitiva é um esforço *contemporâneo* [grifo nosso], com fundamentação empírica, para responder questões epistemológicas de longa data – principalmente aquelas relativas à natureza do conhecimento, seus componentes, suas origens, seu desenvolvimento e seu emprego” (2003, p.19). A descrição de Gardner resume bem sua visão sobre a disciplina, visto que para ele, acima de tudo, a ciência cognitiva traz novas abordagens e respostas para antigos enigmas. Gardner diz ainda (cf. GARDNER, 2003, p.20) que existem cinco principais características que definem essa nova ciência. São elas: 1. a crença de que as

atividades da cognição humana podem ser entendidas como operações sobre representações mentais; 2. o computador como modelo ideal de mente; 3. a desconsideração metodológica dos fatores emocionais e afetivos da cognição; 4. o caráter multidisciplinar; e 5. a preocupação com questões epistemológicas elementares.

Em primeiro lugar, a ciência cognitiva entende que a mente humana, ou melhor, as “atividades cognitivas” são realizadas em nossos cérebros e ocorrem sob a forma de representação. As representações, por sua vez, dizem respeito à forma como os sujeitos enxergam o mundo, isto é, a capacidade da mente em representar fenômenos, eventos e objetos por meio de imagens, conceitos e crenças. A mente é o sistema que implementa e manipula essas representações.

Em segundo, para a ciência cognitiva, o computador é o modelo mais viável de mente. Não podemos esquecer que a revolução cognitiva, embora tenha cortado laços significativos com a metáfora computacional, ainda preserva importantes características, como por exemplo, os paralelos entre os computadores e a mente. A ciência cognitiva, como filha das mudanças decorridas dessas discussões, entende que o computador é um aliado e uma ferramenta indispensável para o estudo da cognição humana.

As três últimas características dadas por Gardner dizem respeito à abordagem metodológica dessa disciplina, em que, se destacam as seguintes características: a ciência cognitiva desconsidera metodologicamente os fatores emocionais e afetivos da cognição porque ainda que sejam importantes fatores, eles não são a prioridade investigativa dessa disciplina. O interesse dos cientistas cognitivos são as funções cognitivas, ou seja, as funções ligadas ao processamento das representações mentais que são a base da aquisição do conhecimento. Sendo assim, fatores emocionais e o conteúdo afetivo são “metodologicamente irrelevantes”. As teorias ou pesquisas cognitivas apresentam (na maioria das vezes) um caráter multidisciplinar. Gardner admite, porém, que entre os cientistas cognitivos existe um desejo de que a ciência cognitiva se torne um campo unificado e independente, no entanto, isso não é uma concessão geral e nos lembra que foi justamente o esforço multidisciplinar que tornou possível o nascimento da disciplina em meados do século XX, e essa continua sendo uma de suas principais características. Já a última característica corresponde à opinião particular do autor sobre a disciplina, a qual, já havíamos comentado. Vejamos melhor agora do que se trata.

Segundo Gardner, existe na ciência cognitiva uma preocupação em lidar com questões

elementares que são muito similares às preocupações da filosofia, em especial, da filosofia clássica. Mas o que levou o autor a estabelecer tais similaridades? Bem, elas surgiram do reconhecimento de que assim como os antigos gregos “[...] os cientistas cognitivos de hoje perguntam o que significa conhecer algo e ter crenças precisas, ou ser ignorante ou estar errado. Eles procuram entender o que é conhecido” (GARDNER, 1996, p.18). E entender o que é conhecido implica justamente em analisar o mundo, os objetos, os sujeitos, os fenômenos, o aprendizado e o raciocínio. Claro, com isso Gardner não quer dizer que a ciência cognitiva está apenas reciclando questões ou retomando o trabalho dos gregos, ao contrário, o autor reconhece que a ciência cognitiva vai além de teorias especulativas, pois o uso da IA e de programas computacionais como parte importante na metodologia de pesquisa são de imprescindíveis para estudo da cognição empreendido pela ciência cognitiva.

No centro das pesquisas que moldaram a ciência cognitiva está um novo programa de pesquisa no campo da psicologia, a saber, a psicologia cognitiva. Segundo Sternberg (cf. 2008, p.25), um dos fatores decisivos para o surgimento do cognitivismo na psicologia foram os primeiros trabalhos em psicobiologia<sup>9</sup>. Esses trabalhos foram liderados por aqueles que estavam insatisfeitos com as explicações behavioristas. No geral, se pode dizer que as críticas ao behaviorismo se deram pela insuficiência do aspecto comportamental para explicar os comportamentos serialmente ordenados (como por exemplo, os movimentos necessários para tocar instrumentos musicais) e pela radicalidade metodológica que tornava inviável o uso de termos “internalistas” como “intenção”, “desejo” e “crença”, visto que eles não podiam ser determinados pela observação comportamental.

O maior problema da teoria behaviorista, então, era a dificuldade de explicar comportamentos complexos visto que muitas dessas explicações necessitavam do uso de termos e eventos internalistas. Um dos primeiros críticos do behaviorismo foi o psicólogo e ex-aluno de Watson, Karl Lashley. A principal insatisfação de Lashley era com o papel “passivo” dado ao cérebro na abordagem behaviorista. A ideia de que o cérebro era um órgão receptor, moldado e influenciado pelo ambiente não lhe parecia certa, pois ele acreditava que o cérebro era determinante e fundamental para o comportamento humano, e isso o fez postular que o cérebro

---

<sup>9</sup> A Psicobiologia investiga a relação entre o desempenho cognitivo (sob determinados eventos) e as estruturas cerebrais. Essa área também é conhecida como “neurociência comportamental”, pois ela utiliza os conceitos da biologia e da neurociência para explicar o comportamento dos seres humanos e também de outros animais.

era, na verdade, o grande organizador (a causa) do comportamento humano. Assim como Lashley, o também psicólogo Donald Hebb criticou o behaviorismo de Watson afirmando que tanto o comportamento quanto a linguagem seriam na verdade determinados pela organização dos conjuntos neurais do cérebro.

Apesar de terem diferentes teorias, as críticas de Lashley e Hebb são similares à medida em que tentam demonstrar o papel determinante do cérebro para o comportamento humano, principalmente, no que diz respeito à linguagem<sup>10</sup>. Essa mudança metodológica (em comparação com a abordagem simbólica) foi especialmente influenciada por novas descobertas da biologia e da neurociência sobre o sistema nervoso de alguns organismos. No livro “Mentes e máquinas”, Teixeira enumera cinco dessas principais descobertas.

a) o comportamento de interceptação para o acasalamento da mosca doméstica; b) o comportamento da rã que projeta sua língua para capturar o inseto que se movimenta à sua frente (estudado por Lettvin, Maturana e McCulloch); c) o comportamento do kingfisher para capturar o peixe dentro da água; d) a distinção entre linhas horizontais e verticais pelo gato recém-nascido; e) o comportamento do cão que abocanha, no ar, a bola atirada pelo dono (TEIXEIRA, 1998, p.122).

Segundo Teixeira, a observação desses comportamentos instigou uma questão fundamental: como esses organismos simples executam comportamentos complexos? Do ponto de vista da abordagem computacional simbólica, a resposta mais aproximada seria que: o comportamento complexo desses organismos é causado por regras computacionais serialmente ordenadas, de modo que, a projeção da língua, no caso das rãs, é o resultado final de uma série de comandos simples, um processo muito similar ao da máquina de Turing.

O problema, no entanto, é que uma descrição computacional desses comportamentos exigiria no mínimo que esses organismos possuíssem um programa complexo como o da mente humana, porém este não é o caso visto que embora possuam uma “consciência inferior” (menos complexa que a mente humana), a mosca, a rã, o peixe e o cão ainda assim são capazes de realizar

---

<sup>10</sup> Outra figura importante é o linguista Noam Chomsky. Na verdade, suas contribuições podem ser consideradas ainda mais significativas que as de Lashley e Hebb, pois o trabalho do autor foi fundamental para tornar a linguística uma disciplina formal, contribuindo significativamente para a revolução cognitiva. Além disso, Noam Chomsky também foi um severo crítico do behaviorismo, em especial, do behaviorismo radical de B. F. Skinner. Dentre muitas alegações, Chomsky definiu a abordagem de Skinner como “limitada” porque ela não explicava o “estado inicial” da linguagem, isto é, ela não conseguia explicar como a linguagem é inicialmente formulada. E para Chomsky, explicar esse aspecto é uma das principais tarefas de qualquer teoria que tenha pretensão de explicar a linguagem.

comportamentos complexos. Essa observação foi o que tornou possível levantar a hipótese de que o sistema nervoso desses animais seria, na verdade, o fator determinante para saber como esses comportamentos eram executados. A consequência imediata dessas observações foi a necessidade de explicar a relação entre o sistema nervoso, os fenômenos mentais e o comportamento.

A preocupação dos psicólogos em entender essas relações foi decisiva para o surgimento do cognitivismo na psicologia, que surgia como uma alternativa ao behaviorismo. Segundo Sternberg, o movimento cognitivista na psicologia surgiu a partir da crença de que grande parte do comportamento humano poderia ser compreendido pelo modo como as pessoas pensam. (cf, STERNBERG, 2008, p. 08). Não obstante, o autor explica que essa nova concepção foi impulsionada pelo uso de simulações por computador e IA que tinham a função de “[...] imitarem uma determinada função ou um determinado processo humano. Entre os exemplos estão o desempenho em tarefas cognitivas específicas (ex.: manipulação de objetos no espaço tridimensional)” (2008, p. 22). Além disso, o uso dos computadores tinha a vantagem clara de explorar uma ampla gama de hipóteses com precisão e praticidade, reforçando a importância do uso dessas máquinas inteligentes no estudo da cognição humana. O cognitivismo, portanto, surge como uma nova forma de entender o comportamento humano com foco no estudo da cognição, ou mais especificamente, nos processos neurobiológicos responsáveis por causar os pensamentos e os comportamentos.

O enfoque nas explicações biológicas sobre a mente não trouxe mudanças apenas a psicologia como também proporcionou uma reavaliação de algumas das questões que moveram a metáfora computacional e os modelos mentais simbólicos. Claro, essas mudanças não ocasionaram uma superação completa da IA simbólica (basta olhar para todos aplicativos e programas que usamos ainda hoje) e tampouco puderam enfraquecer o entusiasmo daqueles que viam no computador um bom modelo de mente, ao contrário, para eles os computadores permaneceram fundamentais porque a linguagem computacional ainda era necessária para caracterizar os processos mentais. O que mudou, segundo Searle (2006, p.285), foi a analogia sobre eles. Se antes a questão essencial era saber se a mente opera como um programa computacional, com a biologia, a questão se tornou “o cérebro é como um computador”. Isto é, levantou-se a hipótese de que os processos cognitivos do cérebro operam de forma semelhante aos de um computador.



Segundo Searle, é justamente essa hipótese que caracteriza o movimento que ele chama de “cognitivismo” e que critica em "A crítica da razão cognitiva". O descontamento do autor com o cognitivismo ocorre por muitas razões, porém, nos adiantemos em dizer que essa nova formulação desagradou a Searle já que ela descreve a mente em termos de estados computacionais sintáticos e segundo o autor desconsidera os aspectos subjetivos e qualitativos da mente.

#### **2.4 Uma nova concepção de mente?**

A metáfora computacional foi um dos maiores marcos para a ciência do pensamento humano, visto que ela iniciou uma série de discussões, abordagens e teorias, como por exemplo, a IA, a ciência cognitiva e o cognitivismo. No entanto, no que diz respeito à filosofia da mente, quais foram de fato as mudanças decorrentes da aproximação entre os computadores e a mente? De acordo com o Professor Cícero Barroso, a mudança mais significativa foi o abandono da concepção de mente baseada na *metáfora do suporte* (cf. BARROSO, 2017, p. 28 e segs).

Desde a filosofia clássica a mente foi compreendida sob o sentido de *lugar*. Uma espécie de local que armazena ou acumula ideias e pensamentos. Essa concepção de mente está inserida nas mais diferentes metodologias e teorias filosóficas e também do senso comum. A ideia de local da metáfora do suporte pode ser apontada por questões simples, como por exemplo, a capacidade humana de reconhecer cheiros, sabores e sons através da memória. Na verdade, somos capazes até mesmo de identificá-los sem os experimentar ou sentir durante longos períodos de tempo. Esse é o caso de memórias antigas, sons e cheiros de infância. Ao memorarmos costumamos sentir diferentes sensações que nos provocam a ideia de nostalgia.

O oposto também pode ocorrer, afinal, nós podemos atribuir cheiros e sons a coisas jamais experimentadas. Uma capacidade que vivenciamos, por exemplo, durante a leitura de livros, jornais ou até mesmo em notícias da televisão. Ambas as situações nos levam a questionar: onde estão esses cheiros, sabores e sons? intuitivamente podemos explicar essas situações com respostas do tipo “guardamos os cheiros, sabores e sons nas memórias dentro da nossa mente” e ou para o segundo caso, dizer “que as pessoas simplesmente imaginam essas coisas em suas mentes”. Ambas as respostas nos levam a perceber que as intuições mais básicas do homem (o senso comum) sobre sua própria consciência parecem sugerir em certa medida, assim como na metáfora do suporte, que a mente é uma espécie de local.

Essa ideia de local baseado na *intuição* do senso comum também se perpetua na filosofia. A tradição racionalista, por exemplo, acredita que a mente é como um local que abriga a razão. É o espaço onde as ideias são formuladas. No empirismo, essa visão não se altera, porém ao invés de ser um lugar onde as ideias são formuladas, a mente dos empiristas é um local de "aprendizado" e armazenamento das informações que obtivemos durante a vida através da experiência. Há também teorias, de ambas as tradições, em que a ideia de local é tratada sob o termo de "substância".

No dualismo, por exemplo, a mente é descrita como "uma substância imaterial que habita o corpo". Platão foi quem primeiro caracterizou essa substância como algo pertencente ao mundo das ideias e somente mais tarde essa substância ressurgiu no dualismo de Descartes como a "res-cogitans". Já na tradição empirista, a mente foi considerada por muito tempo como uma *tábula rasa* que é vazia quando nascemos, mas é preenchida conforme as experiências vividas, criando assim o entendimento e a inteligência dos indivíduos. Para John Locke (o criador da expressão *tábula rasa*), a mente seria uma espécie de "grande armazém" capaz de guardar todas as experiências vividas durante a vida.

Apesar de possuírem diferentes metodologias, todas essas teorias estão de acordo quanto ao fato de que a mente funciona como uma espécie de lugar. E tendo em vista que a metáfora do suporte é utilizada por uma longa tradição de teorias da filosofia da mente, levanta-se a questão: quais são as implicações da substituição da metáfora do suporte pela metáfora computacional?

Primeiramente, a metáfora computacional não aborda os problemas clássicos da mente. Por exemplo, nós sabemos que o pressuposto de que a mente é uma substância diferente do corpo nos leva ao "problema mente-corpo", porém, com a metáfora do computador, esse problema é deixado de lado (e para alguns até mesmo abandonado), pois a concepção de mente da metáfora computacional não é entendida como uma substância e tão pouco é possível supor que a relação "software e hardware" sugira uma relação de oposição como ocorre na metáfora do suporte, visto que a premissa mais básica dos programas computacionais (softwares) reside no fato de que eles precisam do computador (hardware) para serem executados. Dessa forma, não existiriam motivos para se levantar qualquer questão relacionada ao problema mente-corpo.

Outra vantagem do uso da metáfora computacional é que ela não exclui os termos internalistas, pois ao contrário do que pensavam os behavioristas, os cientistas cognitivos não

acreditam que a utilização desses termos possa os conduzir inevitavelmente a algum tipo de dualismo. A razão disso é que os termos internalistas são trabalhados por eles sob a luz da metáfora computacional, que, em certa medida, impede que essas explicações possam indicar uma realidade ou uma substância que seja independente da realidade física. Isso faz com que os termos internalistas não sejam excluídos do vocabulário cognitivista.

Claro, essas vantagens não implicam que a abordagem computacionalista foi capaz de resolver por completo as questões deixadas pelo problema mente-corpo, no entanto, sua metodologia pode ser considerada como “um frescor” ou “alívio”, já que sua proposta sugere deixar de lado as questões decorrentes da concepção tradicional de mente (mente como substância) e se concentrar em compreender, por exemplo, em quais são as relações neurológicas responsáveis pelos fenômenos mentais. Isto é, estudar a cognição humana sem se preocupar com o dualismo.

Apesar de tudo isso, Searle acredita que o cognitivismo está longe de ser verdade. A metáfora computacional e o Computacionalismo, assim como o dualismo, o behaviorismo e o materialismo estão presos ao mesmo problema fundamental: fechar os olhos para o aspecto subjetivo dos fenômenos mentais.

Não há outra razão mais fundamental para a oposição de Searle ao cognitivismo do que esse menosprezo metodológico pelos fenômenos subjetivos e qualitativos. Não por acaso, demonstrar a centralidade da consciência é o objetivo primordial do trabalho de Searle que, segundo ele, permitiria o abandono da dicotomia imposta pelo dualismo e o materialismo, reparando, por consequência, falsas teorias, como o funcionalismo, o Computacionalismo e o Cognitivismo.

Diante de tudo que foi apresentado até aqui, fica claro que o primeiro capítulo teve como principal objetivo expor um quadro geral das teorias e metodologias que segundo Searle são importantes para o contexto do problema da consciência, a começar pelas mudanças ocorridas na ciência moderna. Nosso segundo capítulo, por outro lado, terá dois objetivos fundamentais: apresentar a teoria searliana da consciência e com isso as razões pelas quais o autor a considera o fenômeno mais fundamental relativo à mente, e também introduzir algumas das concepções e argumentos que posteriormente tornarão as críticas de Searle ao cognitivismo mais claras.

### 3 A TEORIA SEARLIANA DA CONSCIÊNCIA

O filósofo americano John Searle possui grandes contribuições na filosofia da linguagem e da mente. No que diz respeito ao campo da linguagem, Searle conta com sua bem-sucedida teoria geral da Intencionalidade e com o êxito em desenvolver posteriormente a teoria dos Atos de fala proposta originalmente por seu mentor, o filósofo John Austin.

Já na filosofia da mente, suas contribuições mais expressivas parecem ser aquelas em que o autor mostrou fortemente seu lado crítico, como é de seu argumento do *Quarto chinês* (um experimento hipotético criado com o objetivo de refutar a Inteligência Artificial Forte) de seu *Princípio da Conexão* (tese contrária à ideia de que existem estados mentais puramente inconscientes) e de sua *A Crítica da Razão Cognitiva* (uma crítica formal ao cognitivismo).

Muito embora essas contribuições tenham sido publicadas em diferentes contextos e com diferentes propósitos, todas possuem uma característica em comum: elas expressam o forte descontentamento do autor para com o tratamento dado aos aspectos subjetivos e qualitativos nas explicações da mente. Isto é, os problemas encontrados pelo autor nessas teorias, é o que torna necessário (ainda que em diferentes proporções) demonstrar a centralidade da consciência para o estudo da mente. No entanto, entender quais são essas razões exige que tenhamos claro, primeiramente, o posicionamento de Searle em relação à consciência. Isto é, devemos entender o que de fato o autor está falando quando se refere à consciência, afinal, tanto o problema evidenciado por ele quanto suas críticas ao cognitivismo se dão fundamentalmente pelo conflito de ideias a respeito do papel da consciência no estudo da mente.

#### 3.1 Por que a consciência é irreduzível?

A afirmação de que certos processos físicos causam as mentes não é um problema para Searle (ou para esta discussão), afinal, ele é um naturalista e conseqüentemente entende que a causa dos fenômenos mentais é inerente aos processos que ocorrem em seu sistema biológico. Mentes, portanto, são causados por processos que ocorrem na estrutura dos cérebros (cf, SEARLE, 2006). Entretanto, as divergências de Searle em relação ao materialismo e ao cognitivismo são provocadas justamente pelas diferentes perspectivas sobre a ideia de redução e subjetividade.

Já falamos sobre a metodologia reducionista na seção 2.2, destacando sua origem e suas vertentes, porém agora, a veremos sob o ponto de vista de Searle. Afinal, porque a metodologia reducionista é um problema? Qual é o posicionamento de Searle sobre o reducionismo tendo em vista sua teoria da mente?

Primeiramente, Searle considera a “irredutibilidade” das mentes uma verdade óbvia. O fato de que a consciência é um fenômeno subjetivo, isto é, uma experiência vivida apenas em primeira pessoa, torna impossível reduzi-la da mesma maneira como tradicionalmente os fenômenos e objetos são reduzidos na ciência (a ideia geral da redução ontológica). O problema dessa alegação é que a crença na irredutibilidade das mentes se contrapõe (em certa medida) à metodologia reducionista e, portanto, pode ser considerada contrária a um método científico consagrado. Como, então, o autor lida com esse conflito? São basicamente dois passos: 1 apontar os problemas da abordagem reducionista; 2 apresentar uma concepção de redução da mente mais adequada.

O primeiro passo então consiste em esclarecer quais são os problemas do reducionismo. Segundo o autor, o problema mais grave do uso das teorias reducionistas reside no fato de que suas exigências metodológicas tornam a discussão da consciência e dos outros fenômenos subjetivos temas quase anti-científicos. Então, no que diz respeito às possibilidades teóricas do estudo da mente, o reducionismo falha porque nos deixa ancorados no mesmo velho dilema em que “[...] os dualistas consideram a irredutibilidade da consciência como prova incontrovertida da verdade do dualismo. E os materialistas insistem que a consciência tem que ser redutível à realidade material” (SEARLE, 2006, p.168). De modo que, a atitude de negar a possibilidade de redução da mente é equivalente a abandonar a ciência e conseqüentemente seguir por um caminho dualista.

Todavia, apesar desse impasse, Searle posiciona-se ao lado dos que negam a redução da mente. Ele acredita que o impasse mencionado seja apenas aparente. Searle acredita que não há porque acreditar que a redução seja a única saída para os problemas dualistas ou que sua negação nos coloque no caminho contrário à ciência. A teoria de Searle é, antes de mais nada, uma tentativa de demonstrar que não há conflitos reais entre a ciência e o estudo dos fenômenos subjetivos da mente. De modo que não há, portanto, a necessidade de recorrer à metodologia reducionista ou qualquer outra que tente radicalmente afirmar a natureza material da mente.

O posicionamento “anti-reducionista” de Searle não é uma novidade trazida por ele para

a filosofia da mente. Pelo contrário, existe uma longa lista de filósofos que são críticos da abordagem reducionista e que se opõem a teorias reducionistas famosas como é o caso da Teoria da Identidade. Ademais, Searle não esconde suas influências teóricas e cita os trabalhos dos filósofos Thomas Nagel, Saul Kripke e Frank Jackson como suas grandes referências anti-reducionistas. O autor também esclarece que, embora no geral esses filósofos possuem teorias que diferem de seu pensamento, ele está de acordo com eles quanto à ineficácia da metodologia reducionista para explicar a mente.

Dentre os nomes citados, Searle destaca com mais ênfase às ideias do filósofo Thomas Nagel, em especial, a tese apresentada em seu famoso artigo "Como é ser um morcego?" de 1974. No artigo em questão, Nagel tenta refutar qualquer possibilidade de redução da mente, alegando que a consciência é na realidade, irreduzível. A tese do autor se tornou uma forte oposição às teorias de identidade, em especial, ao uso das descrições objetivas para caracterizar a mente.

Segundo Nagel, tanto as descrições objetivas quanto as descrições comportamentais behavioristas, não são capazes de descrever em realidade a consciência porque as descrições objetivas são limitadas a sua natureza de terceira pessoa e por isso não podem descrever estados mentais vivenciados/experenciados em primeira pessoa. A descrição objetiva é, e sempre será, uma descrição em terceira pessoa, visto que implica na tentativa de descrever (e nesse sentido de entender) o outro. A descrição objetiva (que é a linguagem científica) é direta e neutra, não podendo haver influência de valores e crenças daqueles que as formulam.

As descrições objetivas, por si mesmas, não são o grande problema aqui, afinal nós as utilizamos rotineiramente no dia-a-dia quando explicamos, por exemplo, que a bola de futebol americana não é redonda ou que o céu muda de cor dependendo da hora do dia. A ciência faz a mesma coisa, porém, utilizando de seu vocabulário próprio. Assim, dados os mesmos exemplos, poderíamos dizer que um cientista ou pesquisador esportivo diria que na realidade a bola de futebol americana possui o formato ovalado ideal para táticas desse esporte. Um cientista também nos explicaria que a cor do céu é vista em azul em decorrência da relação entre os elementos químicos que compõem a atmosfera terrestre, de modo que, o contato do nitrogênio (que compõe 78% da atmosfera terrestre) e do oxigênio (21%) com a luz solar, produz a cor azul, mas dependendo da intensidade da luz, as cores podem variar para tons vermelhos e alaranjados.

Note que por mais diferente ou sofisticado que possa ser o vocabulário empregado nas

descrições objetivas, elas são fundamentalmente descrições diretas sob o ponto de vista de terceira pessoa. Dessa forma, dizer que a bola é ovalada não descreve nada além do seu formato. No entanto, quando precisamos descrever a mente, nossos relatos levam em conta nossa própria experiência mental e, portanto, são descrições subjetivas. Portanto, diferente da descrição objetiva, a descrição subjetiva baseia-se no modo como um indivíduo relata seu ponto de vista sobre o mundo, como ele se refere aos aspectos qualitativos de sua experiência consciente, ou seja, como ele se refere aos qualia.

“Qualia” é o termo utilizado na filosofia da mente para designar os aspectos qualitativos das experiências conscientes. Se trata das propriedades intrínsecas e inseparáveis da vida mental que caracterizam a experiência subjetiva. Eles são o que possibilita que cada mente possua sua visão particular e singular do mundo. Cada indivíduo tem sua própria descrição acerca das experiências vividas, seja por ouvir uma música, provar uma comida, sentir desejo ou apenas olhar para o céu. Todas essas ações serão experienciadas de forma diferente por cada indivíduo, de modo que, mesmo que estejam agrupados e tentem descrever seus estados mentais, a experiência subjetiva seguirá sendo única para cada um daqueles indivíduos.

Nagel explora a questão dos qualia e utiliza a figura do morcego para exemplificar seu ponto instigando o leitor a imaginar como seria ser um morcego. Além de apresentar algumas características únicas do animal, Nagel, menciona alguns de seus comportamentos. Tudo isso para que possamos, ainda que de maneira rudimentar, nos imaginarmos no lugar do mamífero. Em resposta a seu próprio experimento, Nagel relata que:

Até onde eu consiga imaginar isso (e não chego muito longe), isso apenas me diz como seria para mim comportar-me como um morcego se comporta. Mas não é essa a questão. Eu quero saber como é, para um morcego, ser um morcego. Se eu ainda assim tento imaginar isso, fico restrito aos recursos da minha própria mente, inadequados para a tarefa. Não consigo isso nem mesmo imaginando acréscimos à minha experiência presente, nem imaginando segmentos gradualmente subtraídos dela, nem imaginando uma combinação de acréscimos, subtrações e modificações (NAGEL, 2013, p. 250).

Explicar a capacidade do cérebro de produzir estados mentais com aspectos qualitativos é justamente a dificuldade das teorias reducionistas. Afinal, para o autor não há como presumir que descrições objetivas, sejam de caráter neurobiológico, comportamental, ou computacional, possam descrever verdadeiramente os estados mentais conscientes. Assim, Nagel compreende

que as teorias reducionistas em geral estão preocupadas em reduzir toda a vida mental a descrições físicas de suas preferências, sem levar em conta, no entanto, que essas descrições em nada podem descrever de fato como é ser, por exemplo, um morcego, ou como é sentir um cheiro, se sentir apaixonada ou acima de tudo, como é ter uma mente.

O posicionamento de Searle é muito similar ao de Nagel. Ambos concordam que as descrições objetivas são insuficientes para caracterizar a mente. Dessa forma, a redução ontológica da consciência é falha porque existe um conflito entre a experiência em primeira pessoa e as descrições em terceira pessoa. As descrições objetivas são limitadas por sua própria natureza e não podem caracterizar os aspectos qualitativos dos estados mentais e por essa razão, as teorias com modelos mentais que buscam estabelecer essas reduções são enganosas.

Assim, ao invés de recorrer a uma relação de identidade ontológica, Searle propõe que os fenômenos mentais sejam caracterizados em termos de “propriedades emergentes”, isto é, propriedades de nível superior de um sistema complexo S que dependem mas não se reduzem às propriedades de baixo nível de S. Como casos de propriedades emergentes, Searle cita os exemplos da solidez, da liquidez e da transparência.

Searle afirma que a consciência é uma propriedade causalmente emergente de determinados sistemas de neurônios da mesma forma que a solidez e a liquidez são características emergentes de sistemas de moléculas. Da mesma forma, a consciência em si não pode ser deduzida ou presumida a partir da mera estrutura física dos neurônios” [grifo nosso] (SEARLE, 2006, p.162). Ainda que seja causada pelas propriedades de baixo nível do cérebro, ela é uma propriedade inteiramente nova desses sistemas, uma propriedade de natureza subjetiva.

Desse modo, embora descarte a redução ontológica da consciência, Searle propõe ao mesmo tempo um tipo de redução causal da consciência, na qual os poderes causais dos nossos estados conscientes são tão somente os poderes causais do cérebro. Sobre isso, o autor declara: “sustento uma visão de relações mente/cérebro que é uma forma de redução causal conforme defini a noção: características mentais são causadas por processos neurobiológicos” (SEARLE, 2006, p.166). A noção de redução causal não implica, na visão de Searle, uma redução ontológica porque a condição subjetiva (da qual falamos anteriormente) torna a consciência ontologicamente irreduzível. Não há motivos para insistir na redução ontológica, porque o aspecto qualitativo da experiência consciente torna isso impossível.



### 3.2 Consciência: subjetividade, qualitatividade e unidade

Compreendemos a definição searlina de Consciência de duas maneiras, de modo que temos uma pela forma conceitual e outra pela atribuição ontológica. Do ponto de vista formal, é certo afirmar que o autor não oferece uma formulação objetiva clara tendo em vista que Searle não apresenta um conceito único do que seja a consciência. Todavia, os exemplos e os comentários acerca da consciência apresentados em suas principais obras nos fornecem algumas informações para entendermos melhor o que o autor entende por consciência.

Em *A redescoberta da mente*, por exemplo, Searle nos diz que *a consciência é como um interruptor que liga e desliga*. O autor compreende que a mente é um sistema que tem duas formas: ela é consciente ou não. Já na obra *Consciousness and Language*, o autor explica que os estados mentais (aliás, toda a nossa vida consciente) são como “[...] estados de sensibilidade ou ciência que começam quando uma pessoa acorda de manhã, depois de um sono sem sonhos, e se estendem por todo o dia até que ela vá dormir à noite” (2010, p.2). Além disso, como é dito em *Mind, Language and Society*, “a consciência pode acabar quando morremos, entramos em coma ou nos tornamos de outro modo inconscientes” (2000, p.40). A partir das palavras do autor, podemos dizer que em termos formais, a consciência searlina é um fenômeno finito que intrinsecamente é o que torna a mente “desperta”<sup>11</sup>. A consciência faz com que os estados interiores e exteriores sejam percebidos de forma total pelo sistema (mente).

Searle também nos oferece outro tipo de resposta e esta é pautada na questão ontológica. De acordo com essa resposta, “[...] a consciência não é uma “coisa”; é uma característica ou propriedade do cérebro no sentido em que, por exemplo, a liquidez é uma característica da água” (2006, p.155). A consciência, portanto, “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” (2006, p.133). Em *Minds, Brains and Science*, Searle volta a reafirmar que “[...] todos os fenômenos mentais, quer conscientes ou inconscientes, visuais ou auditivos, dores, cócegas, comichões, pensamentos, na realidade, toda a nossa vida mental, são causados por processos que têm lugar no cérebro” (2017, p. 23). Esses processos são

---

<sup>11</sup> Searle usa os termos “acordar” e “despertar” como exemplos para estados de consciência. Em contraposição, os termos “desligar”, “dormir” e “morrer” são exemplos de estados de inconsciência.

neurobiológicos pois ocorrem na estrutura do cérebro onde os neurônios, as células, as sinapses e o sistema nervoso se relacionam.

Ademais, outro traço importante na teoria de Searle é o modo como ele busca identificar e esclarecer as principais características da consciência. Dentre aquelas que o autor compreende serem as mais fundamentais estão: a subjetividade, a qualitatividade e a unidade. A subjetividade é o aspecto que torna a consciência única e tão diferente dos outros fenômenos biológicos do corpo humano (esse é precisamente o aspecto que a torna, para muitos, um objeto inadequado para o estudo científico). O fato de a consciência ser subjetiva torna sua existência uma experiência única e exclusiva de cada organismo consciente. Os estados mentais também possuem um aspecto que é qualitativo. Searle explica que “[...] os estados conscientes são qualitativos no sentido de que há, para cada um deles, um determinado modo de senti-lo, uma determinada característica qualitativa” (SEARLE, 2000, p. 42). Isto não vai muito além do que já vimos na discussão anterior sobre os qualia. Estados mentais são qualitativos e não há maneira de separar ou ignorar este fato.

Há também o fato de que os estados mentais conscientes atuam sob forma de unidade. A visão searliana de consciência é contrária à abordagem comum, na qual a consciência é entendida sob a ideia de “blocos de construção” em que as peças são dispostas isoladamente (são fragmentadas), mas se encaixam para formar uma figura completa. Já Searle compreende que a maneira mais correta de entender a consciência é pela abordagem do “campo unificado”, em que os estados subjetivos ocorrem simultaneamente e unificadamente. Não partes, mas um todo.

Em suma, a consciência é biológica e não fragmentada. Ela não ocorre de forma isolada, pelo contrário, ela existe sob a forma de unidade, na qual, todos os estados mentais são experienciados, sentidos, e capturados pelo cérebro de forma integral. É a maneira como os fenômenos ocorrem de maneira global e unificada (ou holística<sup>12</sup>). Não existe uma forma sequencial ou lógica nos fenômenos mentais. Eles não estão dispostos em uma fita à espera de comandos. Uma pessoa pode, por exemplo, sentir dores nos ombros, fome, cansaço, calor e outros diferentes estados mentais que embora tenham intensidades diferentes ocorrem ao mesmo tempo.

Searle também explica que a mente, de maneira geral, atua como um “grande

---

<sup>12</sup> A visão holística (ou o holismo) é um conceito das ciências naturais em que um sistema é visto como um todo. Essa visão se opõe à ideia de olhar para os fenômenos e as coisas como sendo a soma de suas partes.

organizador” do conjunto de relações dos organismos com o meio ambiente e também seus próprios estados:

Através das modalidades sensoriais, por exemplo, o organismo obtém informações conscientes sobre o estado do mundo. Ouve sons nos seus arredores; vê objetos e estados de coisas em seu campo de visão, percebe odores específicos de aspectos distintos de seu ambiente etc. Além de sua experiência sensorial consciente, o organismo terá também, caracteristicamente, experiências de ação. [...] podemos dizer que na percepção consciente o organismo tem representações causadas por estados de coisas no mundo, e, no caso de ações intencionais, o organismo provoca estados de coisas no mundo através de suas representações conscientes. (SEARLE, 2006, p. 157-158)

Embora as três características mencionadas anteriormente sejam consideradas pelo autor como fundamentais, a intencionalidade é central para caracterizar os estados e fenômenos mentais, a razão disso, como veremos a seguir é que a intencionalidade é a capacidade das mentes de “se dirigirem” para algo além de si mesmas. Claro, há também outros aspectos da consciência, Searle as categoriza como “o campo da consciência”, “a estrutura Gestáltica”, “a familiaridade” e “o humor”. Esses termos representam a “estrutura da consciência”. Algumas delas já foram mencionadas, mas receberão uma apresentação formal com as devidas considerações do autor.

### 3.3 Intencionalidade

A intencionalidade é responsável por fazer a “ponte” entre o interior de nossas mentes e o mundo externo, é que a propriedade de um estado mental que possibilita que ele represente o mundo, que ele seja sobre algo. Trata-se de uma característica intrínseca dos estados mentais, ou seja, algo que depende unicamente da natureza da mente<sup>13</sup>.

Historicamente, a intencionalidade remonta ao período medieval, sendo um tema tradicionalmente discutido pelos fenomenólogos. O filósofo Franz Brentano foi quem tornou a intencionalidade essencial para caracterizar a consciência, afirmando que ela é a característica

---

<sup>13</sup> É interessante esclarecer que a Intencionalidade não é uma capacidade restrita aos seres humanos. Searle acredita, diferente do que se estabeleceu tradicionalmente (principalmente pela forte herança racionalista), que mentes humanas não são especiais e embora elas possuam um alto grau de Intencionalidade, isso não significa dizer que os animais não tenham intencionalidade. Existem, por exemplo, algumas espécies de mamíferos, como por exemplo, os chimpanzés, elefantes e golfinhos que possuem mentes que apresentam significativa inteligência, consciência e intencionalidade. As diferenças entre as mentes desses mamíferos e a mente humana não está em uma categoria ontológica, mas é consequência do processo natural de evolução das espécies. Essa diferenciação, no entanto, é mais evidente sob o aspecto linguístico.

que diferencia os fenômenos psicológicos dos fenômenos físicos. Posteriormente, a intencionalidade também foi discutida por Edmund Husserl, que definiu a intencionalidade como a capacidade da mente de “se dirigir” para algo que está para além de si mesma. Diferente de Brentano, Husserl acreditava que não havia necessidade de classificar os estados mentais como intencionais ou não-intencionais, pois a consciência era sempre intencional.

As discussões sobre intencionalidade são extremamente importantes na filosofia de Searle, pois essa temática, além de lhe render prestígio acadêmico, foi de certa maneira, a porta de entrada do autor no estudo filosófico da mente. Em sua obra *Intentionality*, publicada em 1983, o autor não apresentou severas mudanças conceituais em relação ao Husserl já havia estabelecido, no entanto, Searle foi quem caracterizou e transformou a intencionalidade em uma teoria geral com temas em comum entre a filosofia da linguagem e da mente.

Sobre a Intencionalidade, Searle explica através de exemplos práticos que “[...] se eu tiver uma crença, deve ser uma crença de que determinada coisa é desse ou daquele modo; se tiver um temor, deve ser um temor de alguma coisa ou de algum acontecimento; se tiver um desejo, deve ser um desejo de fazer alguma coisa, ou de que algo aconteça” (SEARLE, 2002, p.1). Os estados mentais existem em respeito a algo, alguém ou por algum acontecimento que pode ter acontecido ou não, pois o autor considera que os estados intencionais de imaginação também são formas de intencionalidade.

Não se pode esquecer que a imaginação também se baseia em representações mentais. Por isso, mesmo em casos onde a imaginação está completamente errada, como por exemplo, acreditar que pinguins são capazes de voar, ainda assim, ela será intencional. Pois o conteúdo mental do estado “pinguins podem voar” faz de alguma forma referência ao mundo. Ora, pinguins também possuem asas assim como outras aves, então imaginar que eles poderiam voar não seria algo tão fora da realidade, embora seja verdade que pinguins não possam, de fato, voar.

Perceba, no entanto, que o esforço de Searle é apenas de caracterizar a Intencionalidade. A validade dos estados mentais não é a questão fundamental aqui, mas sim o conteúdo representacional que eles empregam. Pois como explica Searle, a “Intencionalidade é aquela propriedade de muitos estados e eventos mentais pela qual estes são dirigidos para, ou acerca de objetos e estados de coisas no mundo” (SEARLE, 2002, p.1). Dessa forma, o conceito de intencionalidade descrita pelo autor é pura e simplesmente a capacidade das mentes de representar coisas, propriedades e fenômenos.

Além disso, a compreensão completa do conceito de Intencionalidade requer o entendimento de outras duas categorias subjacentes. São elas: a intencionalidade *intrínseca* e a intencionalidade *como-se* que correspondem respectivamente a duas formas de imprimir a intencionalidade. A primeira delas é considerada por Searle como a intencionalidade real e verdadeira e a segunda como “ilusória”, pois a intencionalidade do tipo como-se é uma forma de atribuição mental que os sujeitos concedem às coisas e os fenômenos.

Para exemplificar seu ponto, Searle pede para pensarmos em um rio. A água que escorre pelo rio pode ser descrita em termos de intencionalidade. Por exemplo, um sujeito X poderia dizer que as correntes estão tentando chegar em determinado lugar e por isso elas seguem rápidas em determinados trechos. Esta seria uma análise decorrente da observação de X do percurso da água. Searle, no entanto, acredita que validar esse tipo de atribuição de intencionalidade é muito perigoso, pois isso nos colocaria em uma situação em que não haveria como negar, por exemplo, que tudo é mental, afinal a intencionalidade como-se depende exclusivamente da forma como algo parece intencional aos olhos de seus observadores.

Por outro lado, a intencionalidade intrínseca não é algo fruto de atribuição. Ao contrário, ela é intencional por natureza (intrínseca ao seu sistema físico). Tanto os estados mentais inconscientes quanto os conscientes são formas de intencionalidade intrínseca. No caso dos estados conscientes, ninguém dirá, por exemplo, que tem uma crença ou que odeia, ama e acredita de forma metafórica. Ao contrário, esse tipo de estado é afirmado de forma literal. Os sujeitos creem, amam e acreditam. São estados mentais reais.

Da mesma forma, não há como dizer que alguém tem sentimentos reprimidos de forma metafórica, ou que existem razões inconscientes pelas quais um sujeito não gosta de demonstrar afeto ou carinho. Essas afirmações, embora falem do inconsciente, também não são afirmações reais. Pois existem estados mentais que não são intencionais, como é o caso da ansiedade ou da depressão que surge sem uma razão específica e que não pode ser atribuída a fatores externos. Claro, tanto a ansiedade quanto a depressão podem ser causadas por razões além do sujeito, mas essa não é uma regra universal. Nesses casos, não existe uma explicação para esses estados porque eles não dizem respeito a algo ou alguém, isto é, eles não têm um objeto ou um fator que os impulsiona.

Por último, devemos falar sobre as sensações corpóreas. A interação entre o corpo e a mente é manifestada não apenas pelos cinco sentidos (visão, tato, olfato, paladar e audição), mas

também pela consciência corpórea, como é o caso, por exemplo, da consciência da posição dos membros (braços e pernas) em relação ao corpo e da sensibilidade dos dentes. De acordo com Searle, todos esses estados sensoriais podem ser “agradáveis e desagradáveis”. Dores são desagradáveis e o cheiro de rosas é agradável, por exemplo.

Há também outras características e condições da intencionalidade que Searle considera importantes<sup>14</sup>, no entanto, são temas secundários no que diz respeito aos fins propostos por este trabalho, sendo assim eles não serão abordados, embora algumas questões ainda sejam retomadas posteriormente.

### **3.4 O Campo da Consciência e seus aspectos**

Apesar da consciência atuar como uma unidade, Searle diz que existe um campo de atenção que determina como as coisas são percebidas. Uma situação comum é a de alguém que se machuca, como por exemplo, um corte nos dedos, mas ela sente fome ou tristeza antes mesmo de notar a dor ou sangramento (ou vice-versa). Isso ocorre, segundo Searle, porque o campo da consciência se forma pelo centro e pela periferia e o que está no centro da consciência se sobrepõe ao que está na periferia. No centro estão os estados aos quais os sujeitos estão prestando atenção no momento. Dessa forma, seguindo o exemplo, o sujeito está ciente do corte nos desde e perifericamente, está sentindo fome, mas somente depois de ver um letreiro de um restaurante é que ele se dá conta disso. A atenção (o centro) do sujeito se torna a sua sensação de fome e a dor nos pés torna-se secundária

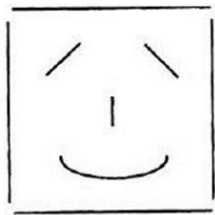
Claro, isso é um exemplo simplista. O mais correto é que o sujeito esteja sentindo diversas outras sensações ou vontades ao mesmo tempo. Ele pode se arrepender de usar o sapato apertado ou ficar impaciente com a demora do garçom em lhe trazer um cardápio. A quantidade de sensações, estados e desejos é irrelevante nessa questão. O que queremos focar é a forma como Searle entende o funcionamento da consciência. Sendo assim, o mais importante é esclarecer que o centro e a periferia não são formas sequenciais ou lógica dos estados, mas que existem níveis de atenção que se alternam por questões qualitativas ligadas à percepção, os

---

<sup>14</sup> Ficam de fora as discussões sobre os mecanismos da intencionalidade como é o caso do conteúdo proposicional e da direção de ajuste. Para uma visão completa da teoria de Intencionalidade de Searle, recomenda-se o livro *Intentionality: An Essay in the Philosophy of Mind* de 1983.

sentidos e a intencionalidade. Há também mais três aspectos: a estrutura gestáltica, a familiaridade e o humor.

Outro ponto importante sobre os estados mentais é o fato de que eles atuam sob uma forma de estrutura gestáltica da experiência consciente. A estrutura gestáltica é uma teoria psicológica de que a percepção trabalha as representações sob a organização de certas estruturas. Ou seja, os estímulos visuais são organizados em um padrão que se assemelha a formas familiares do nosso campo perceptivo. Para entender melhor, tomemos a figura abaixo:



(Figura original da obra: SEARLE, J. *Mente, linguagem e sociedade: filosofia no mundo real*. Rio de Janeiro: Rocco, 2000)

Embora a figura apresentada não seja de fato um rosto humano, nossos cérebros através do estímulo visual externo de nossa percepção captura as linhas e traços de modo que os vejamos como um rosto humano. A Gestalt trata justamente disso, é sobre a maneira como a percepção apreende e interpreta os estímulos visuais. A hipótese de Searle, no entanto, é de que essa percepção nada mais é do que nossa capacidade intencional consciente e não apenas o resultado da captação visual. Essa forma estrutural depende também de outro aspecto e Searle o chama de familiaridade.

A característica anterior está ligada ao que Searle chama de aspecto da familiaridade, que por sua vez consiste na capacidade da mente se relacionar com as coisas, sujeitos e fenômenos por familiaridade. Esse aspecto é o que garante que a "bola quadrada" desejada pelo personagem Quico do seriado Chaves, seja completamente inimaginável, muito embora bolas de futebol não sejam quadradas. Os estados de familiaridade são um fenômeno que surge por graus maiores ou menores. Além disso, a familiaridade é intrínseca aos estudos mentais. E assim como outros aspectos da mente, a familiaridade só é rompida em casos de patologia, como ocorre por exemplo, com a síndrome de Capgras<sup>15</sup> em que o sujeito não se reconhece por familiaridade.

---

<sup>15</sup> A síndrome de Capgras é um transtorno psicológico que desenvolve a crença delirante de que uma pessoa foi substituída por um "impostor", ou seja, uma pessoa que é psicologicamente diferente, mas fisicamente idêntica ao substituído.

Além das formas estruturais e da familiaridade, o humor também é considerado por Searle como uma característica fundamental da consciência. O humor mencionado pelo autor diz respeito ao "estado de espírito" do sujeito consciente. Ele é, nesse sentido, as vivências afetivas do indivíduo que podem ir da mais forte euforia até a mais profunda tristeza.

Claro, não há humor apenas em sentimentos extremos e os indivíduos também são acometidos de momentos em que não estão felizes ou tristes. Há tédio, chateação e preocupação também. O processo de escrever uma dissertação, por exemplo, pode deixar um sujeito angustiado e preocupado por alguns meses até finalmente sentir alívio e felicidade no momento da defesa. As mudanças, bem como a própria causa do humor, podem se dar pelos mais diversos estados mentais, bem como por fenômenos extremos.

### 3.5 A Inconsciência e o Princípio da Conexão

Searle acredita que desde a teoria do inconsciente de Freud, passamos a classificar os estados mentais inconscientes como sendo mais importantes do que os conscientes, principalmente, no estudo da cognição humana como é o caso da percepção, memória e linguagem. Essa inversão de valores se deu pela falta de clareza de algumas noções conceituais, sobretudo, da noção “ingênua” do que sejam os estados mentais inconscientes. Não por acaso, o autor propõe, no capítulo sete da obra *A redescoberta da mente*, soluções a esses problemas na tentativa de tornar as relações entre a consciência e o inconsciente mais claras.

A teoria do Inconsciente apresentada pelo neurologista e psiquiatra Sigmund Freud<sup>16</sup> em *Traumdeutung*<sup>17</sup> nos anos de 1900, repercutiu no mundo todo (dentre muitas razões) por defender que o inconsciente não era algo físico escondido no interior dos cérebros (algo que era fortemente difundido na época), mas um fenômeno psíquico manifestado nas mentes. Uma forma simples e resumida para exemplificar essa relação pode ser dada pelo seguinte esquema:

Estado mental (+) **Consciência** = Estado mental consciente

Estado mental (-) **Consciência** = Estado mental inconsciente

<sup>16</sup> As informações sobre Sigmund Freud são baseadas, sobretudo, em comentários de autores e de livros que falam da temática. Para essa seção não investigamos a fundo as obras do autor.

<sup>17</sup> A obra é conhecida no Brasil pelo título em português “A Interpretação dos Sonhos”.



Sendo assim, um estado mental inconsciente é um estado mental, porém, com ausência da consciência. Searle explica que esse tipo de abordagem faz um forte apelo à ideia de que seria possível “extrair” a consciência dos estados mentais e ainda assim teríamos um estado mental, embora inconsciente. Ou seja, se retirarmos a consciência o que restará é o inconsciente. Uma consequência dessa visão é que os estados mentais inconscientes parecem atuar, segundo Searle, como uma espécie de parasita que habita nos estados mentais conscientes. Alguns estados estão infectados e outros não.

Searle discorda de tudo o que foi dito até aqui e por mais que ele reconheça o trabalho e a importância das contribuições de Freud, acredita que essa noção de inconsciente tem que ser esclarecida já que ela é prejudicial para a compreensão do real papel da consciência na vida mental, sobretudo, para que tenhamos condições de resolver os problemas tradicionais da filosofia da mente. Não por acaso, o autor propõe que a consciência é na verdade a característica fundamental para determinar aquilo que de fato é mental.

Essa afirmação é defendida pelo autor através do argumento do Princípio da Conexão, no qual, Searle afirma que “não temos nenhuma noção do inconsciente, a não ser como aquilo que é potencialmente consciente” (SEARLE, 2006, p.218). Ou seja, o que designamos como “mental” somente pode ser entendido como de fato “mental” em decorrência de sua relação com a consciência. A hipótese defendida pelo princípio da conexão é então de que os estados mentais e qualquer outro fenômeno mental somente é percebido pelos sujeitos em termos de consciência.

Assim, admitido que a mentalidade é algo atribuído aos estados por meio da consciência, a teoria de Searle acaba por criar uma grave consequência: como explicar os estados mentais inconscientes? Ora, se os estados dependem de sua relação com a consciência para serem caracterizados como mentais, como é possível que haja estados mentais que são inconscientes?

Para Searle isso não é possível. Não existem estados que são mentais e também puramente inconscientes. Essa afirmação é dada pelo autor como duas finalidades: primeiro negar a existência de estados que sejam mentais e ao mesmo tempo inconscientes e em segundo, firmar a consciência como sendo o fenômeno fundamental para determinar aquilo que é ou não mental. Ambas as pretensões são fundamentais para a tese searliana de consciência, todavia elas carecem, antes de mais nada, uma explicação convincente sobre qual seria o papel do inconsciente, ou melhor, qual é a relação entre o inconsciente e o consciente.

Para o autor, entender essa relação requer esclarecer aquilo que pode ser considerado de fato inconsciente, afinal, o que existe na mente que pode ser considerado inconsciente? Searle acredita que há uma clara diferença entre o grupo de estados mentais que são de fato inconscientes e aqueles que são apenas não conscientes. Essa diferença é precisamente descrita como:

Considero a diferença, por exemplo, entre minha crença (quando não estou pensando sobre isso) de que a torre Eiffel está em Paris e a mielinização dos axônios em meu sistema nervoso central. Há um sentido em que ambas são inconscientes. Porém, há uma diferença entre elas no sentido de que os estados estruturais de meus axônios não poderiam ser em si estados conscientes, porque não há nada mental neles" (SEARLE, 2006, p. 221- 222).

A passagem revela que embora a mielinização<sup>18</sup> seja importante para a realização dos estados mentais, o autor considera que seu papel nos processos causais que produzem esses estados não é suficiente para caracterizá-la como um estado propriamente mental, seja ele consciente ou inconsciente, já que mielinização e todos os processos neurofisiológicos não apresentam formas de intencionalidade ou quaisquer outros aspectos mentais.

Não obstante, Searle acredita que se alguém fosse capaz de sentir a movimentação dos axiomas em seu cérebro, ainda assim, essa movimentação não poderia ser considerada estados mentais reais porque a única coisa que o sujeito poderia conscientemente saber sobre ela, estaria restrito à sua experiência (sensação) sentida em seu cérebro.

Com o exemplo dos axônios, Searle pretende estabelecer duas asserções importantes: em primeiro lugar, existem as crenças inconscientes, como por exemplo, a crença de que a torre Eiffel está em Paris, e existe também a atividade neuronal necessária para a realização desse estado mental. Em segundo lugar, está o entendimento de que ambos (as crenças e a atividade neuronal) ocorrem no cérebro, porém, apenas a crença inconsciente pode ser considerada verdadeiramente mental. Dessa forma, os fenômenos de mielinização são considerados formas "não conscientes" e os estados que apresentam intencionalidade, como as crenças e os desejos, são estados inconscientes. A distinção entre os estados mentais inconscientes e os não conscientes é determinada justamente por sua relação com a consciência.

---

<sup>18</sup> A mielinização corresponde ao processo feito pela camada mielina na proteção dos neurônios e a condução de informações até os axônios dos neurônios. Sem esse processo, a transmissão de informação neuronal seria lenta.

Diferente das formas “não conscientes” (mielinização), as crenças inconscientes (estados inconscientes) possuem relação (conexão) com a consciência por preservarem intencionalidade em sua forma. E ainda que na maior parte do tempo sejam crenças inconscientes, esses estados podem ser conscientemente "acessados" em algum momento, isto é, a presença de intencionalidade torna esses estados potencialmente acessíveis à consciência. Com isso, os estados mentais inconscientes e os não conscientes diferem-se pelo conteúdo mental e sua relação com a consciência.

A diferença entre esses estados se torna mais clara quando Searle os caracteriza em termos de intencionalidade. Para isso, é preciso rapidamente lembrar que existem duas formas de classificar os conteúdos intencionais dos estados mentais. São eles: a intencionalidade intrínseca e a como-se, em que, de maneira geral, podemos dizer que a intencionalidade intrínseca é aquela que é própria dos estados mentais e a como-se é uma intencionalidade por atribuição feita pelos sujeitos. Todos os estados mentais conscientes são formas de intencionalidade intrínseca, no entanto, Searle acredita que os estados inconscientes também podem ser considerados dessa forma. Por exemplo, se alguém tem uma crença ou um desejo de forma inconsciente, esse desejo ou essa crença, existe de forma intrínseca, isto é, o conteúdo intencional desses estados ocorre de forma objetiva não dependendo de observadores ou sujeitos além daquele que detém esses estados. Ou seja, os estados inconscientes que são intencionais são intrínsecos assim como os estados mentais conscientes.

Por outro lado, também existem estados em que a intencionalidade ocorre apenas por atribuição como ocorre em casos de intencionalidade como-se. Para Searle, os processos neurofisiológicos e a mielinização se encaixam nesses casos, visto que a noção de que esses processos são de alguma forma intencional é adotada exclusivamente em função dos sujeitos. Em outras palavras, o autor está dizendo que a única forma de intencionalidade de estados não conscientes é atribuída como-se tivessem. Ademais, além da classificação em intencionalidade intrínseca e como-se, devemos lembrar que o conteúdo mental intencional dos estados conscientes e inconscientes também possuem uma certa forma aspectual.

O termo forma aspectual designa uma característica universal da intencionalidade. Essa característica, por sua vez, é aquilo que faz dos estados mentais formas de experiência, isto é, a maneira como um fenômeno ou objeto é apreendido. Por exemplo, “[...] quando você vê um carro não é simplesmente uma questão de objeto sendo registrado por seu aparato perceptivo;

em vez disso, você realmente tem uma experiência consciente do objeto de um certo ponto de vista e com certas características” (SEARLE, 1990, p. 587). Essas características aspectuais da experiência são, portanto, aquilo que tornam um estado mental intencional e isso pode ser verificado tanto em estados mentais conscientes quanto inconscientes.

Além disso, as formas aspectuais possuem sua condição própria e são limitadas às experiências particulares dos sujeitos. Por isso, elas não podem ser caracterizadas em termos de terceira pessoa ou por descrições comportamentais. Elas também não podem ser definidas por descrições computacionais ou processos neurofisiológicos como na mielinização. Searle diz que “haverá sempre um abismo inferencial entre os motivos comportamentais epistêmicos para a presença do aspecto e a ontologia do aspecto mesmo” (SEARLE, 2006, p.227). Isso ocorre porque essas descrições, por mais precisas que sejam, não imprimem o caráter aspectual dos estados intencionais das experiências vividas.

Todavia, é importante esclarecer que isso não significa que os estados e suas formas aspectuais são independentes. Não há vida mental sem os processos neurofisiológicos que ocorrem no cérebro, afinal, a mente é um fenômeno biológico. No entanto, esses processos, por si só, não são suficientes para caracterizar as formas aspectuais de intencionalidade dos estados e experiências vividas. Neste ponto, chegamos à parte mais importante do princípio da conexão e ela toca precisamente o seguinte problema:

Parece que temos uma contradição: a ontologia da intencionalidade inconsciente consiste inteiramente em fenômenos neurofisiológicos, objetivos, de terceira pessoa, mas ainda assim os estados têm uma forma aspectual que não pode ser constituída por tais fatos, porque não há nenhuma forma aspectual no nível de neurônios e sinapses (SEARLE, 2006, p.229)

Dessa forma, a questão a ser respondida é: Como os processos neurofisiológicos podem formar ontologicamente os estados mentais inconscientes, mas não serem suficientes para descrever ou constituir tais estados? A resposta a esta questão foi mencionada na introdução desta seção quando dissemos que “a inconsciência é um fenômeno acessível à consciência da mesma forma como ocorrem com os outros fenômenos mentais” e a evidência está no fato de que, apesar de serem inconscientes, esses estados apresentam formas aspectuais de intencionalidade.

Como também já sabemos, Searle acredita que a intencionalidade e a consciência não

podem ser dissociadas, pois a intencionalidade é manifestada nas mentes em termos de consciência. Entretanto, se esse é o caso, como isso pode ocorrer com os casos de estados inconscientes? Searle acredita que a explicação decorre do fato de que os estados inconscientes são formas de possibilidade da consciência, isto é, estados inconscientes podem (ainda que nunca ocorra) se tornarem conscientes, pois seu conteúdo intencional e sua forma aspectual conecta esses estados “adormecidos” à consciência. Em outras palavras, o conteúdo intencional intrínseco dos estados inconscientes os torna em certa medida ou em algum grau acessíveis à consciência.

Claro, existem também os casos de estados conscientes não intencionais, como por exemplo, a exaltação, a depressão, ansiedade etc. Esses casos são conscientes, mas não são intencionais, pois não há nenhum um objeto intencional e, portanto, não podem ser considerados estados intencionais. Para entender melhor essa relação Searle afirmar que:

Embora nem todos os estados conscientes sejam intencionais, e nem todos os estados intencionais sejam conscientes, existe uma ligação essencial: só compreendemos a intencionalidade em termos de consciência. Há muitos estados intencionais que não são conscientes, mas são o tipo de coisa que poderia ser potencialmente consciente (SEARLE, 2009, p.36).

Searle mais uma vez quer chamar atenção ao fato de que tanto os estados mentais conscientes quanto os inconscientes, sejam eles intencionais ou não intencionais, existem enquanto “conteúdos” potencialmente acessíveis à consciência. Somente em termos de consciência é que podemos atribuir mentalidade aos estados e aos fenômenos em nossas mentes.

Searle também propõe aplicar o princípio da conexão em casos de “dor inconsciente” em que temos a seguinte situação: “pessoas que sofrem de dores crônicas, digamos dores crônicas nas costas, que em decorrência delas às vezes fique difícil adormecer. E na verdade, depois que já adormeceram, às vezes há ocasiões, durante a noite, em que seu estado de saúde precário faz com que elas acordem” (SEARLE, 2006, p.236). Como o Princípio da conexão pode explicar esses casos? Primeiramente, Searle ressalta que não podemos ignorar que a dor é real. Ela ocorre ainda que a pessoa esteja em completo estado de inconsciência (dormindo), porém podemos assumir, por exemplo, que a dor esteve consciente em algum momento e em outro momento esteve inconsciente. Searle explica que o que permite que o estado de dor permaneça é o processo neurofisiológico desse estado de dor que agora está inconsciente, mas que em algum

momento esteve consciente.

Em suma, no que diz respeito à discussão sobre o inconsciente, a teoria de Searle se apresenta da seguinte maneira: toda a vida mental é ontologicamente possível graças aos processos neurobiológicos que ocorrem no cérebro. Os fenômenos mentais produzidos por esses processos podem ser conscientes ou inconscientes. Os inconscientes podem ser classificados como mentais ou não mentais. A diferença entre eles é dada pelo conteúdo intencional intrínseco. Aqueles que possuem esse conteúdo são capazes de causar estados conscientes, já os que não possuem não podem causar estados conscientes. Ou como explica o autor:

Não há nada acontecendo em meu cérebro a não ser processos neurofisiológicos, alguns conscientes, alguns inconscientes. Dentre os processos neurofisiológicos inconscientes, alguns são mentais, outros não. A diferença entre eles não está na consciência, porque, por hipótese, nenhum é consciente; a diferença é que os processos mentais são candidatos à consciência, porque são capazes de causar estados conscientes. Mas isso é tudo. (SEARLE, 2006, p.232)

Tendo essas alegações em mente, Searle conclui que não existem estados mentais "puramente inconscientes", isto é, estados mentais não conscientes, pois um estado mental pressupõe a existência de um conteúdo mental intrínseco. Algo que os processos neurobiológicos não conscientes não podem determinar sem recorrer à consciência e aos aspectos subjetivos. Aceitar o princípio da conexão significa aceitar também que não existe um inconsciente, pelo menos não da forma como tradicionalmente ele foi estabelecido. Também é aceitar que não existem estados mentais puramente inconscientes, ou no caso do Computacionalismo e do Cognitivismo, que haja "sentenças dentro da cabeça" das quais o sujeito permanece completamente inconsciente. Se Searle estiver correto, não apenas as teorias advindas da metáfora computacional seriam refutadas, mas também as metodologias reducionistas. Afinal, não haveria redução de fenômenos e estados conscientes a fenômenos e estados puramente inconscientes.

### **3.6 E agora? O que fazer com a consciência?**

De fato, desde a introdução deste trabalho temos falado sobre as questões em torno da consciência, em especial, do conflito gerado por sua natureza subjetiva e a objetividade das

ciências. Neste ponto, porém, cabe questionar: como Searle acredita ser possível tornar a consciência um objeto de estudo das ciências naturais? Para Searle, o primeiro passo deve ser a resolução ou melhor, o esclarecimento de algumas falácias e os erros associados à noção de objetividade no âmbito científico.

Sabemos que o problema da objetividade científica no âmbito do estudo da mente ocorre porque a natureza da consciência e de outros aspectos mentais é subjetiva e a relação entre a subjetividade e a objetividade é naturalmente conflitante. No entanto, no que diz respeito à mente, Searle acredita que é “um erro assumir que a subjetividade das mentes se opõe ao princípio da objetividade científica” (2000, p.38), pois essa oposição à subjetividade das mentes é na verdade fruto de um erro interpretativo das noções de objetividade epistemológica e objetividade ontológica.

A objetividade epistemológica é usada como um critério necessário para a validade de enunciados, teorias e hipóteses. Por sua vez, a objetividade ontológica é sobre atribuir realidades físicas à natureza das coisas. O problema, no entanto, é que os dois estão constantemente ligados um ao outro, de modo que, acredita-se que o objetivo último da ciência deva ser a atribuição de uma realidade exclusivamente material (objetiva) para todas as coisas existentes.

Searle, por outro lado, discorda dessa afirmação e diz que a ciência tem que se preocupar em ser epistemologicamente objetiva, afinal, como foi descrito pelos positivistas, na ciência “[...] aspira-se à limpeza e à clareza, recusam-se distâncias obscuras e profundezas insondáveis” (NEURATH, 1986, p.10). O dever científico é antes de mais nada a produção de conhecimentos objetivos acerca do mundo, pois a verdade com a qual a ciência está comprometida é (e deve ser) concreta e assertiva. Entretanto, o autor nos alerta que é imprescindível não confundir as questões epistemológicas com as ontológicas. Em razão disso, o autor diz que:

Aquilo em que quero insistir incessantemente é que podemos aceitar os fatos óbvios da física - por exemplo, que o mundo é constituído inteiramente de partículas físicas em campos de força - sem, ao mesmo tempo, negar os fatos óbvios de nossas próprias experiências - por exemplo, que somos todos conscientes e que nossos estados conscientes têm propriedades fenomenológicas irreduzíveis bastante específicas. O erro é supor que essas duas teses são incompatíveis. (SEARLE, 2006, p.44-45)

Claro, boa parte da natureza, as plantas, os objetos, seres e fenômenos são

ontologicamente objetivos, no entanto, isso não significa dizer que não existem fenômenos de natureza subjetiva, afinal, a consciência e todas as formas de estados e fenômenos mentais são a prova disso. Descartes estava certo; não podemos negar nossa racionalidade (mente), pois as experiências subjetivas são reais. Ora, se alguém sofre um acidente grave, a experiência subjetiva da dor que se sucede é tão real quanto às lesões pelo corpo. E ainda que a natureza da consciência seja subjetiva isso não impede que ela seja estudada e investigada cientificamente, pois a pretensão da ciência é a objetividade epistemológica e não há motivos para acreditar que ela se contraponha à subjetividade ontológica das mentes



## 4 AS CRÍTICAS DE SEARLE AO COGNITIVISMO E AS OBJEÇÕES QUE ELAS ENFRENTAM

O capítulo anterior foi escrito em favor de dois objetivos principais: em primeiro lugar, havia a necessidade de apresentar a concepção searlina de consciência, e em segundo, a de esclarecer as razões que fazem Searle concluir que a consciência é central para caracterizar a mente. Já neste capítulo, abordaremos as críticas do autor aos modelos de mente computacionais, mais especificamente, a Inteligência Artificial Forte e o Cognitívismo, levando em conta sua teoria da consciência. Em sequência, discutiremos também algumas das objeções feitas às alegações de Searle com o foco central na sua crítica ao cognitívismo e a seu princípio da conexão, visto que ele é fundamental tanto para as objeções de Searle ao cognitívismo quanto para sua concepção de mente.

### 4.1 As limitações da Metáfora Computacional

Na Filosofia, Searle é considerado um dos mais importantes críticos do Computacionalismo. Esse status foi concedido a ele, principalmente, por seu famoso argumento do Quarto chinês da década de 1980. Com o Quarto Chinês o autor pretendia refutar a Inteligência artificial simbólica sob o argumento de que tanto o uso da sintaxe para caracterizar os estados mentais quanto o uso do Teste de Turing como validação da inteligência dos computadores não poderiam garantir (nem mesmo perto disso) uma simulação da mente humana natural.

Evidentemente, Searle não é o único crítico do Computacionalismo ou da metáfora computacional. Nomes como de Hubert Dreyfus e Thomas Nagel são alguns das importantes peças que completam o quadro geral de objeções à visão computacionalista de mente. Apesar das diferenças teóricas, Dreyfus, Nagel e Searle estão de acordo quanto ao fato de que existem severas limitações nas explicações computacionais, principalmente, sobre o papel causal dado aos fenômenos qualitativos. Vejamos brevemente alguns de seus comentários.

Em 1972, Hubert Dreyfus publicou a obra *What computers can't do?* que reunia importantes objeções à Inteligência Artificial simbólica, através da comparação entre o

modelo representacional de mente<sup>19</sup> em contraposição à visão fenomenológica das faculdades humanas de entendimento e compreensão. O modelo representacional mencionado por Dreyfus é o computacional, mais especificamente, aquele instaurado inicialmente por Alan Turing e difundido pelos engenheiros e matemáticos do instituto de tecnologia de Massachusetts<sup>20</sup>.

Em contraposição, ao modelo representa simbólico (e em virtude de sua sólida base na fenomenologia de Martin Heidegger), Dreyfus (assim como Searle) parte do princípio de que os seres humanos compreendem e significam as coisas de forma holística, isto é, a compreensão de algo é sempre a compreensão integral e total do seu fenômeno. Mas o que isso significa exatamente? Que as coisas (objetos, fenômenos e seres), ao contrário de como é sugerido pelo modelo representacional de Turing, não são percebidas pelos sujeitos de forma única e sequencial (como os comandos da fita), mas sim de forma integral pela mente e por todos os sentidos.

Dreyfus explica ainda que situações do dia-a-dia exigem dos sujeitos decisões rápidas que são baseadas em nosso “sistema de relevância”. Esse sistema exige que as informações a serem usadas sejam “filtradas” por “relevância”. Esse tipo de situação (comum aos seres humanos) exigiria do programa de computador uma mudança na maneira como os outputs (dados a serem gerados) e inputs (dados coletados) deveriam ser processados.

O “senso de relevância” na escolha das informações é algo primordial no entendimento e na linguagem humana. Essa característica, no entanto, não pode ser encontrada nos modelos representacionais de mente porque eles são limitados a sua forma atomística em que a compreensão é resultado de uma síntese formada da combinação de dados isolados (inputs) que são processados de forma sequencial. Assim, nos modelos representacionais, os dados são predicados exclusivamente pela forma que já foi previamente estabelecida no programa, não atendendo assim um dos requisitos básicos da linguagem e do processamento de informações realizados pelo homem. O contraste entre o modo de compreensão humana e o processamento de informação dos computadores fez Dreyfus concluir que o resultado previsto no jogo da imitação de Turing não passaria de uma grande “utopia”, já que os computadores não seriam

---

<sup>19</sup> Na verdade, a ideia de “representação mental” remonta ao filósofo Platão e posteriormente a grandes nomes da modernidade, como por exemplo, René Descartes, Thomas Hobbes e John Locke. Esses autores, utilizam o termo “ideia” como representações mentais. No entanto, o modelo mencionado no trabalho de Dreyfus é o da IA simbólica.

<sup>20</sup> O modelo é como a máquina de Turing que funciona a partir de um conjunto de regras formais encarregadas de processar informação sob a manipulação de símbolos.

capazes de simular a compreensão humana num teste real em que se necessitasse do senso de relevância para responder às questões lançadas. As contribuições do filósofo Thomas Nagel também foram importantes objeções ao Computacionalismo, principalmente, por chamar atenção para o problema da redução das experiências subjetivas, ou melhor, para o problema da irreduzibilidade das mentes. Nós já vimos os argumentos que levam o autor a negar a possibilidade de redução das mentes (veja o tópico 3.1 sobre a irreduzibilidade da mente”), mas é importante esclarecer para o autor o problema persiste no Computacionalismo e IA, visto que ambos requerem aceitar que tanto os estados quanto às experiências subjetivas são passíveis de reduções a partir de descrições computacionais. E já sabemos que Nagel acredita que isso é impossível visto que a descrição objetiva (seja ela computacional ou não) é limitada por sua natureza de terceira pessoa, o que por consequência, torna inviável acreditar que os computadores possam experimentar o mundo como ocorre nas mentes naturais.

Em suma, enquanto Dreyfus chama atenção para a visão holística, Nagel evidencia uma dificuldade latente do uso metodológico das descrições objetivas. Assim, embora distintas, ambas as críticas apontam para um caminho em comum, no qual, o modelo computacional se mostra insuficiente para caracterizar a mente humana. Searle também apresentou seu ponto de vista sobre os problemas do Computacionalismo, destacando sobretudo as limitações da sintaxe computacional para caracterizar a linguagem das mentes. Essa discussão foi o que deu origem ao já mencionado, argumento do Quarto Chinês. Vejamos agora como o autor trabalha essas questões.

#### **4.1.1 O Quarto Chinês**

Como foi dito ainda na introdução deste trabalho, Searle é um crítico severo da ideia de que o computador pode ser um bom modelo mental. Ele rejeita tanto as comparações em que a mente é descrita como uma espécie de programa computacional (como é o caso da IA simbólica) quanto as que acreditam que cérebro é um tipo de computador digital (como no cognitivismo). Esse tópico é sobre a IA simbólica, ou melhor, a IA Forte (nome utilizado por Searle nas obras cf. 1990; 2006; 2017). O termo “forte” corresponde à categorização dada por Searle às vertentes da IA, em que ele as classifica em dois grupos: a IA do tipo Fraca e a IA do tipo Forte.

De modo geral, a diferença entre essas duas concepções reside no fato de que a IA fraca

não atribui estados mentais aos programas computacionais como ocorre na concepção forte. Searle explica (cf, 2017, p.37) que a IA de sentido fraco entende que o computador é um importante instrumento de pesquisa que permite testar hipóteses de uma maneira mais precisa. Além disso, o computador também pode realizar tarefas complexas e com isso simular inteligência. Já no sentido forte, acredita-se que, se adequadamente programado, um computador pode ser considerado uma mente artificial. Tudo depende apenas de encontrar o programa de mente certo.

Searle não faz objeções a IA fraca, pois considera a concepção Forte fundamentalmente mais problemática e as razões para isso são muitas, porém, as mais patentes são a aceitação de que os estados mentais são resultados de pura manipulação simbólica e ainda que a vida mental consciente é exclusivamente computacional.

Ademais, Searle diz que aceitar esse tipo de tese requer aceitar também duas sérias conclusões: a primeira é de que o cérebro (o sistema físico) é irrelevante, pois se a mente pode ser entendida simplesmente por meio de seu “programa”, o material, a forma e o tipo de “hardware” se tornam de pouca importância. Ou seja, os cérebros são triviais em comparação às mentes. Em segundo lugar, está o entendimento de que as crenças e desejos podem ser experienciados por qualquer tipo de máquinas, como por exemplo, as calculadoras, afinal, elas só precisam que sejam implantados os inputs e outputs corretos. Em termos formais baseados nos princípios Computacionalistas, não há nada de errado em assumir tais alegações.

Obviamente, Searle as considera absurdas e contrariamente afirma que os programas computacionais e sua linguagem sintática não são suficientes para caracterizar ou determinar de forma satisfatória a mente biológica. Para o autor, existe na linguagem computacional uma limitação que é latente e que diz respeito à questão semântica. Não por acaso, o objetivo do Quarto Chinês (QC) é justamente demonstrar que as diferenças entre a linguagem computacional e a linguagem humana é o que torna impossível assumir que um programa de computador possa ter uma mente. Vejamos então, sem mais delongas, o que ocorre no experimento<sup>21</sup>:

---

<sup>21</sup> 13 O experimento do QC apresentado é uma forma resumida da apresentação original feita no texto: LIMA, S. N. S. Searle e a inteligência artificial: uma crítica a partir do Naturalismo Biológico. Monografia (graduação) – Universidade Estadual do Piauí - UESPI, Curso Licenciatura Plena em Filosofia, Campus Prof. Alexandre Alves de Oliveira, Parnaíba-PI, 2020.

<b>Premissas</b>	<p>Imagine alguém sozinho em um quarto que está vazio exceto por uma cesta com símbolos chineses e um manual de instruções na língua natural do sujeito. No manual existem descrições de como organizar e manipular os símbolos chineses da cesta.</p> <p>Imagine também que no exterior do quarto existe uma outra pessoa e que ela está encarregada de enviar algumas perguntas em chinês por debaixo da porta.</p> <p>Apesar da pessoa dentro do quarto não entender chinês, ao seguir as instruções do manual, ela consegue organizar os símbolos e formular respostas satisfatórias às perguntas enviadas.</p>
<b>Conclusão</b>	<p>Por receber respostas às suas perguntas, a pessoa no exterior do quarto conclui que o sujeito no quarto compreende perfeitamente a língua chinesa.</p>

Agora, se substituísimos a pessoa por um programa de computador, como por exemplo, a máquina de Turing, Searle diz que é provável que vejamos duas falsidades nas alegações do Computacionalismo. Em primeiro lugar, está o fato de que o teste de Turing não é um método confiável para estabelecer ou validar a inteligência dos computadores. Pois da perspectiva do QC é possível notar que embora a pessoa no quarto (nesse caso, o programa de computador) tenha formulado respostas corretas aos olhos de quem está de fora (especialista do teste), a pessoa dentro do quarto ainda permanece sem compreender a língua chinesa. Ou seja, a máquina passaria no teste de Turing, no entanto, isso não quer dizer que ela de fato compreende o significado dos símbolos. O seu “aparente” conhecimento em mandarim é na verdade reflexo de sua eficiência em reproduzir a organização dos símbolos contida nas instruções (isto é, nas regras de sua programação).

A segunda alegação é sobre a insuficiência da linguagem sintática para estabelecer a compreensão. Para Searle, não é possível supor que as operações computacionais baseadas em termos de pura sintaxe são suficientes para afirmar que os computadores possuem estados mentais ou que são capazes de compreender coisa alguma. A linguagem computacional é sintática, isto significa dizer que ela consiste apenas de símbolos e regras para combinação de símbolos em sentenças simples e de sentenças simples em sentenças complexas, sem a adição de

nenhum tipo de significado.

Assim, o problema abordado se resume em: os programas são sintáticos, mas a sintaxe não é suficiente para garantir semântica. As mentes são intrinsecamente semânticas, logo implementar um programa é insuficiente para que haja mente. Dessa forma, o QC demonstra que se a compreensão dos significados dos termos é irrelevante para os programas é impossível supor que a sintaxe por si só possa gerar processos semânticos. Isso independe da complexidade do programa. Mesmo um programa altamente complexo como o Chat GPT, em última análise, seria apenas um mecanismo gerador de estruturas sintáticas, e, portanto, ainda seria como o quarto chinês. Por certo, a IA do Chat GPT possui uma linguagem natural extremamente avançada, incluindo funções como tradução, produção de textos e conversação em alto nível, mas, ainda assim, suas capacidades são apenas o resultado de manipulação simbólica.

#### ***4.1.2 A Crítica da Razão Cognitivista***

Após o sucesso do Quarto Chinês na década de 1980, Searle concluiu que havia refutado de forma definitiva a IA Forte, no entanto, decidiu prosseguir com suas críticas ao Computacionalismo em palestras, livros e artigos posteriores. Porém com as mudanças nas discussões sobre o Computacionalismo e a revolução cognitiva, o autor passou a desaprovar as postulações da Ciência Cognitiva e do Cognitivism. A questão subjacente era: O que há de fato nos computadores que os tornam semelhantes às mentes?

Em 1990, com o artigo *Consciousness, explanatory inversion, and cognitive science*, Searle definiu o foco central de suas objeções ao afirmar que o maior erro das concepções advindas do Computacionalismo era a negligência da centralidade da consciência no estudo da mente e que isso tornava as teorias cognitivistas cheias de erros conceituais e metodológicos. Além do artigo de 1990, Searle dedicou também o texto *The Critique of Cognitive Reason* (a Crítica da Razão Cognitivista) a análise de algumas das principais alegações do Cognitivism com foco na hipótese computacionista de que os cérebros são uma espécie de computador digital.

A crítica de Searle fundamenta-se sobretudo no discurso de que existem “três grandes dificuldades” na abordagem cognitivista. São elas: a ausência de uma definição clara do que é Computação; o erro categorial ontológico do conceito de sintaxe (na qual, Searle diz que ela não

pode ser considerada uma característica intrínseca da física) e por último a falácia de que os cérebros processam informações.

### **Dificuldade I - A Definição de Computação**

Para a Ciência Cognitiva, a consciência e as variadas formas de estados mentais são entendidas como resultados de processos computacionais que ocorrem no nível mais baixo. Esse nível é compreendido pelos cientistas cognitivos como descrições sintáticas. Dessa forma, eles compreendem que os estados mentais são, em algum grau ou a princípio, estados sintáticos que se relacionam a partir de uma lógica interna que é anterior a qualquer tipo de fenômeno mental. Ou seja, estados mentais são causados por estados computacionais inconscientes. Eles são como sentenças dentro da cabeça.

Searle acredita que essa definição, embora simples, exige que os leitores tomem como verdadeiras algumas ideias implícitas, como é o caso, do princípio de que o cérebro de fato possui um nível computacional, ou melhor, da suposição cognitivista de que os estados mentais são de fato computacionais. Mas o que exatamente essa suposição quer dizer? O que é essa computação que supostamente existe nos cérebros?

Para essas questões, Searle categoricamente responde: não existe uma definição clara do que seja o computador ou computação e por essa razão é difícil supor que esses conceitos caracterizam qualquer sistema físico, em especial, as mentes. Para o autor é no mínimo curioso o fato da mente ser constantemente comparada ao computador, sendo que nem sequer existe um consenso universal ou um conceito formal para o que seja o computador.

A falta de consenso, segundo Searle, também torna os fundamentos clássicos postulados por Alan Turing (como a ideia de manipulação simbólica) fundamentais e relevantes para a computação ainda na contemporaneidade. Outra questão também postulada por Turing é a ideia de que os componentes físicos das máquinas são irrelevantes. Há um entendimento de que qualquer coisa que possua um programa pode ser considerada um computador. Em suma, uma vez que um componente físico X pode rodar um programa, o material do sistema pouco importará. Dessa forma, se entende que um computador pode ser feito de qualquer coisa. Claro, essas definições não são um consenso universal, mas Searle insiste que elas são fundamentais em grande parte dos trabalhos de computação, principalmente, pela influência de Turing na criação dessa disciplina.

Esses dois fundamentos, no entanto, dão a Searle uma perspectiva desanimadora, pois se considerarmos válidas as características citadas, nossa conclusão não poderia ser outro senão de que os cérebros são como computadores e que estados mentais poderiam ser realizados nos mais diversos materiais desde que o sistema pudesse ser programado adequadamente. Em outras palavras, concluiremos que a mente poderia ser reproduzida por qualquer coisa.

Searle brinca quanto a essa questão e diz que em razão disso alguém poderia construir um cérebro de “gatos e ratos” ou de “canos de água”. Obviamente isto parece impossível, no entanto o autor acredita que não é formalmente incorreto se as alegações de Turing forem tomadas como premissas. Em suma, Searle acredita que falta uma melhor definição para o que seja um computador e também porque as definições de Turing são incompletas, abre-se longas margens para conclusões e hipóteses absurdas como cérebros feitos de gatos.

### **Dificuldade II - A sintaxe não é intrínseca à física**

As teorias computacionalistas possuem uma intrincada relação com a Sintaxe, afinal ela é a característica que define a linguagem dos computadores e o conceito de computação clássica. No QC, entretanto, vimos que a manipulação simbólica não é suficiente para que um programa possa de fato entender. No contexto do Cognitivismo, porém, a sintaxe surge como característica ontológica do nível causal dos estados mentais. Contrariamente, à alegação de Searle é de que a sintaxe não é uma propriedade intrínseca da física e, portanto, não pode caracterizar níveis reais de qualquer sistema físico.

De maneira geral, o cognitivismo compreende que os estados mentais existem em níveis de descrições. Existem os níveis mais altos (isto é, pensamentos e os fenômenos mentais) e os de nível inferior que são os mais básicos. No nível básico se assume que o cérebro trabalha com alguma forma sintática de descrição, como por exemplo, a linguagem binária dos computadores. Os dois níveis correspondem a estados mentais, no entanto, todos os níveis acima do básico não são considerados reais, pois os níveis mais altos são conseqüências do processo que ocorre no nível básico. Dessa forma, tentar descrevê-las ou concentrar discussões nesses aspectos é algo desnecessário, afinal o único nível real é o nível sintático (computacional). Admitir esse tipo de afirmação, no entanto, nos leva a dois problemas fundamentais: a questão da múltipla realizabilidade e a falsa noção de sintaxe.

O argumento da múltipla realizabilidade sugere que um mesmo estado mental X (uma



propriedade, fenômeno ou evento) pode ser realizado por diferentes sistemas físicos. Ora, se os estados mentais são definidos como estados computacionais, que por sua vez, são descrições puramente sintáticas, a sintaxe é, portanto, independentemente de qualquer sistema físico.

Mas existe um segundo problema: A sintaxe não é intrínseca à física. Searle explica que embora a linguagem computacional seja definida em certos termos e símbolos, a sua existência é condicional. Computadores operam em um nível simbólico, mas não há nada físico que se possa encontrar em sua estrutura que possa ser considerado sintático. A questão é simples: a atribuição de propriedades sintáticas é sempre relativa a um agente ou observador que trata certos fenômenos físicos como sintáticos. (SEARLE, 1992, p. 841). Se a sintaxe é admitida em termos do observador, dizer que os cérebros são como computadores consiste simplesmente em submeter o cérebro a nossa própria noção de computação.

Além disso, a realizabilidade múltipla de processos computacionalmente equivalentes em meios físicos diferentes não é apenas um sinal de que os processos são abstratos, mas de que eles não são, de modo algum, intrínsecos ao sistema. Eles dependem de uma interpretação de fora. (SEARLE, 1992, p.842). Dessa forma, a despeito da proposta Cognitivista, a questão que se coloca é: Há algo nos cérebros que os torna intrinsecamente um computador? Searle categoricamente diz que não, ao menos não enquanto a definição de computador for baseada na concepção sintática de Turing. O que não apenas nega a possibilidade de que haja correlação entre computadores e mentes como também dificulta a resolução do problema de realizabilidade universal. Reverter esse quadro exige que a definição de computação se torne mais rígida e que de fato possa nos permitir estabelecer equivalências físicas reais entre os computadores e as mentes.

### **Dificuldade III: O cérebro não processa informações**

Searle, nos avisa que sua visão a respeito desse tópico é um tanto quanto polêmica e que ele está ciente de que muitos o criticarão, mas acredita ser extremamente necessário expor seu ponto de vista porque envolve uma questão fundamental, a saber, a noção que temos sobre processamento de informações. Devemos lembrar que as correlações entre computadores e mentes se iniciaram por uma afirmação básica: os computadores são semelhantes à mente à medida que sua principal característica é o processamento de informação.

Na verdade, os primeiros modelos computacionais, como por exemplo, a máquina de Turing e os sistemas simbólicos que precederam o ENIAC, foram inicialmente projetados através

da aproximação do processo que a mente realiza ao resolver um cálculo; a ideia era justamente operar sob uma lógica que imitasse a inteligência humana. Sendo assim, a ideia de que mentes e computadores processam informação de forma semelhante nunca foi de fato questionada (ao menos essa é a posição do autor). Se este é o caso, então para Searle faz muito sentido que a metáfora computacional tenha sido tão popular, afinal, para entender a mente basta entender quais procedimentos os computadores devem executar para que seja possível processar informações da mesma forma que os cérebros.

Todavia, Searle acredita que essa visão é equivocada. Computadores não processam informação como os cérebros porque na realidade os cérebros não processam informação. Searle diz que isso se caracteriza como “falácia do homúnculo”, pois consiste em “tratar o cérebro como se houvesse algum agente dentro dele, usando-o para efetuar computação” (SEARLE, 2006, 303). Para que possamos ver que isso é um erro, o autor pede que tentemos descrever o que os computadores fazem e os comparemos ao cérebro. Em uma simples avaliação, Searle diz que é possível observar que o computador depende exclusivamente de sua programação, e esta depende de um agente externo encarregado de codificar regras que determinam como as informações são processadas pelos componentes do computador.

Os processos determinados pelo programa, como visto, não são intrinsecamente sintáticos, e por isso os outputs obtidos pelo computador são considerados sintáticos ou semânticos apenas como forma condicionada pelo observador. A parte física também não importa, já que um computador pode ser criado por diferentes materiais e seu sucesso em realizar determinado processamento ocorre sob responsabilidade do programa.

Por outro lado, os processos neurobiológicos que ocorrem no cérebro são processos reais que não dependem ou muito menos são relativos ao observador. Há, claro, as descrições que fazemos dos fenômenos. Essas descrições são construídas a partir de um observador ou ponto de vista, no entanto, o que é condicionado pela descrição não determina aquilo que ocorre no interior dos cérebros. O ponto debatido por Searle é então de que o sentido que temos de “processamento da informação”, tanto científico quanto do senso comum, estão em um nível alto de abstração que tornam os computadores ou qualquer outro sistema de processamento de informação incapaz de “capturar o concreto”. Para esclarecer seu ponto, Searle apresenta o seguinte exemplo:

Suponha que eu veja um carro vindo em minha direção. Um modelo

computacional padrão de visão receberá informações sobre a matriz visual na minha retina e, eventualmente, imprimirá a frase: "Há um carro vindo em minha direção". Mas isso não é o que acontece na biologia. A realidade biológica não é a de um monte de palavras ou símbolos sendo produzidos pelo sistema visual; ao contrário, trata-se de uma questão concreta e específica de um evento visual consciente. (SEARLE, 1992, p. 844-845)

Assim, entende-se que embora o computador possa realizar de forma impecável o processamento de informação de determinados eventos e fenômenos, ainda assim restará o fato de que os fenômenos não são informações a serem processadas formalmente, mas estados de experiência vivenciados pelos sujeitos. A negação do papel da experiência subjetiva, na visão de Searle, é o maior erro que qualquer teoria da mente poderia cometer, pois a experiência subjetiva, ou melhor, a subjetividade é aquilo que essencialmente caracteriza as mentes. Em suma, podemos dizer que a negligência com a consciência e os aspectos subjetivos é o que mais desagradava a Searle nos modelos de mente cognitivistas.

De um modo geral, podemos resumir a discussão das duas últimas seções sublinhando que os programas de computador são definidos em termos de sintaxe e as mentes, por outro lado, são semânticas. Por oposição, a mente não é um programa de computador. Já a hipótese cognitivista entende que a mente é um computador digital, na medida em que suas estruturas cognitivas são regidas por uma lógica interna muito semelhante ao computador. Mas Searle considera um erro caracterizar as operações do cérebro em termo de sintaxe computacional porque a sintaxe não é uma característica inerente da física e por isso não pode ser atribuída a sistemas físicos.

#### **4.2 Existem problemas nas alegações de Searle contra o Cognitismo?**

Segundo o psicólogo e também filósofo Gustavo Arja Castañon as objeções de Searle apresentadas no Quarto chinês são muito assertivas e parecem refutar de forma definitiva o Computacionalismo clássico (do tipo de Turing) e também a IA do tipo Forte. Entretanto, ele acredita que o mesmo não pode ser dito das críticas de Searle ao cognitismo porque o autor comete dois erros fundamentais:

##### **I. Searle “confunde” o cognitismo com o Computacionalismo;**

## II. Searle acusa o cognitivismo de não reconhecer a importância da consciência para o estudo da cognição humana.

Dadas a discussão proposta por Searle em *A crítica da razão cognitiva*, Castañon está certo em afirmar (I), pois a definição de cognitivismo apresentada pelo autor (2006; 1990) não caracteriza a proposta metodológica cognitivista dos psicólogos cognitivos. Searle chama de cognitivismo “a visão de que o cérebro é um computador digital” (cf, 2006, p. 288) e suas explicações posteriores se intercalam para reforçar que essa é de fato a ideia central que move o cognitivismo (seção 4.1.2). Ideias estas que ele diz serem fundamentada em várias especulações:

Se alguém examina os livros e artigos que sustentam o cognitivismo, encontra determinadas assunções comuns, frequentemente não enunciadas, mas não obstante disseminadas [...] também se admite que a questão sobre o fato de os processos cerebrais serem ou não computacionais é simplesmente uma questão empírica e evidente. Isto deve ser estabelecido por investigação factual, da mesma forma que questões como o fato de o coração ser uma bomba ou de as folhas verdes realizarem a fotossíntese foram estabelecidas como questões de fato (SEARLE, 2006, p.291)

É sobre essa visão que Searle insere “as grandes dificuldades do cognitivismo”. A questão que Castañon insere, porém, é de que o cognitivo ilustrado por Searle parece muito mais uma versão atualizada da IAF e do Computacionalismo do que de fato o cognitivismo dos psicólogos cognitivos. Na verdade, Castañon afirma que “Searle não leva em consideração em sua obra a distinção entre o cognitivismo como movimento psicológico, as ciências cognitivas como empreendimento multidisciplinar e a inteligência artificial, tomando todos como expressões do computacionalismo” (CASTAÑON, 2006, p.104). O autor também explica que o erro cometido por Searle é bastante comum entres as críticas humanistas<sup>22</sup> do cognitivismo. Em geral, eles parecem não se preocupar em distinguir termos como “ciência cognitiva”, “cognitivismo” ou “psicologia cognitiva”.

Para enfatizar seu ponto, Castañon explica que o cognitivismo é na verdade um movimento teórico que nasceu da psicologia cognitiva, ou seja, é uma “doutrina” adotada pelos

---

<sup>22</sup> 13 O autor chama de críticas humanistas aquelas oriundas da fenomenologia, filosofia e das teorias humanistas experimentais americanas. O autor também explica que as críticas humanistas são englobadas por “identidade de perspectivas – e adesão às teses básicas humanistas de realidade irreduzível do fenômeno da consciência, defesa da existência de um nível real de liberdade subjetiva e da especificidade da inteligência humana em relação a todas as outras formas de inteligência conhecidas” (CASTAÑON, 2006, p.285).

psicólogos cognitivos (seção 2.3.1). A principal diferença é que a *psicologia cognitiva* “se refere à área de pesquisa específica. Enquanto área de pesquisa, a psicologia cognitiva pode se definir como o estudo de como seres humanos percebem, processam, codificam, estocam, recuperam e utilizam informação. É o estudo do processamento humano de informações” (CASTAÑÓN, 2006, p.97). Além disso, o autor explica que o cognitivismo é marcado por cinco características fundamentais, as quais, Penna ilustrou:

A primeira é a centralidade do conceito de regra para explicar o processamento cognitivo e o comportamento. A segunda o comprometimento com uma visão construtivista dos processos cognitivos. A terceira pela concepção do comportamento humano como orientado a metas. A quarta é a imagem de um sujeito ativo, e não reativo como o da tradição positivista. Por fim, a quinta seria a recuperação do conceito de consciência na psicologia. (PENNA, 1984 apud CASTAÑÓN, 2006, p.97)

Enquanto isso, o Computacionalismo em sua formulação clássica (o autor se refere a Turing, mas especialmente a Fodor) define-se por ser “a crença de que os processos mentais consistem essencialmente em manipulação simbólica lógica, ou seja, computação. Também é conhecida pela abreviação CTM (computational theory of mind), e é associada classicamente a Jerry Fodor e seu clássico "Language of Thought", de 1975” (CASTAÑÓN, 2006, p.97). Como vimos, o Computacionalismo de Fodor institui que nossos estados mentais são como “estados simbólicos processados de forma sequencial” de forma que a linguagem da mente (ou linguagem do pensamento) é o resultado das relações e transformações lógicas dos símbolos.

Em razão disso, Castañon acredita que quando Searle, por exemplo, sugere que o cognitivismo inevitavelmente nos leva a uma múltipla realizabilidade (de que é possível criar uma mente de gatos e ratos porque a linguagem computacional de 0's e 1's pode ser aplicada a qualquer coisa) ou quando ele diz que o cognitivismo é uma espécie de “falácia do homúnculo” (um agente dentro do cérebro, cf: seção 3.5), ele não está afetando diretamente os cognitivistas, mas que definitivamente, Searle acerta em cheio as postulações dos computacionalistas.

Também há o caso (II) para o qual Castañon cita os nomes dos psicólogos

estadunidenses George Miller e Jerome Bruner<sup>23</sup> como dois representantes do resgate do conceito de consciência para o cognitivismo. Bruner, em especial, defendia o conceito de intencionalidade da consciência e a importância do significado para a cognição, de modo contrário às ideias do Computacionalismo clássico. Segundo Castañon, Bruner acreditava que “o ser humano [...] é um foco de atividade do universo. Busca ativamente metas, constrói ativamente suas estruturas cognitivas, atribui ativamente significado” (CASTAÑON, 2006, p.106). Além de Miller e Bruner, Castañon também cita os nomes de mais dois psicólogos estadunidenses, Ulrich Neisser e Roger Sperry<sup>24</sup>, que apresentam teorias em que a consciência é vista como uma propriedade emergente. Castañon também diz que elas possuem similaridades à teoria de Searle (cf, CASTAÑON, 2006, p.107).

Os casos de Miller, Bruner, Neisser e Sperry evidenciam que o cognitivismo, diferentemente do Computacionalismo clássico e da IAF, não pode ser identificado à tese de que o mental pode ser reduzido à manipulação simbólica, ou seja, não pode ser identificado ao computacionalismo. Dessa forma, para Castañon, se aplicarmos *A crítica da razão cognitiva* à última proposta mencionada, não há dúvidas de que as objeções de Searle são fundamentais, mas se elas forem direcionadas ao cognitivismo, o autor está certo de que as implicações (I) e (II) devem ser seriamente consideradas como equívocos da análise de Searle.

### 4.3 O problema da Forma Aspectual

Por certo, a maior parte da argumentação searliana em favor da centralidade da consciência é reservada ao Princípio da Conexão (PC), na qual, a concepção de “mentalidade” é algo que só pode ser estabelecido por meio exclusivo da consciência. No entanto, concordar com o PC leva a algumas consequências, como por exemplo, aceitar o fato de que não existem fenômenos mentais puramente inconscientes, afinal, da perspectiva de Searle, não é possível que um estado seja mental e ao mesmo tempo totalmente inconsciente.

Afirmar que o “status” de mentalidade depende da relação desses estados com a

---

<sup>23</sup> As informações sobre os dois autores são baseadas na discussão proposta por Castañon, por essa razão deixaremos aqui a referências das obras indicadas pelo autor: MILLER, G. (1985). *The Constitutive Problem of Psychology*. In: Koch, S. e Leary, D. (orgs). *A Century of Psychology as Science*. (pp. 40-46). New York: McGraw-Hill Book Company. BRUNER, J. (1997). *Atos de Significação*. Porto Alegre: Artes Médicas.

<sup>24</sup> NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts. SPERRY, R. (1993). *The Impact and Promise of the Cognitive Revolution*. *Am. Psychologist*, 48, 878-885

consciência é tanto a base da teoria de Searle como das suas insatisfações com as metodologias computacionalistas, visto que a metodologia baseada na síntese e manipulação simbólica torna os processos mentais puramente inconscientes. Dessa forma, a explicação de Searle sobre os estados mentais inconscientes é fundamental não apenas para suportar sua teoria geral, mas também para suas críticas ao cognitivismo. Entretanto, são justamente as declarações de Searle sobre o inconsciente o ponto mais debatido entre seus críticos. Uma objeção em especial, diz que as explicações Searlianas sobre os estados mentais inconscientes são insatisfatórias porque o autor não esclarece de fato qual é a real força causal desses estados para a produção de fenômenos mentais. Por consequência, questiona-se a centralidade da consciência para a mentalidade.

Como explica Tárík de Athayde Prata<sup>25</sup>, a concepção disposicional nas explicações de Searle sobre os fenômenos mentais inconscientes são, em certa medida, o calcanhar de Aquiles de sua teoria da consciência. Isso se deve ao fato de que para o autor “ao mesmo tempo em que concebe os fenômenos inconscientes como disposições, Searle atribui a eles algum tipo de eficácia causal” (PRATA, 2017, p. 390). Ora, se for possível concluir, a partir das explicações searlianas, que os estados mentais inconscientes são capazes de intervir causalmente, a centralidade da consciência é fragilizada e inviabilizaria por sua vez, o princípio da conexão.

Ademais, se Prata estiver correto, não apenas o princípio da conexão pode ser colocado em questão, mas também suas críticas ao cognitivismo, afinal, parte das objeções de Searle são construídas sob o entendimento de que não existem fenômenos mentais puramente inconscientes. Sendo assim, se o problema está nas explicações searlianas sobre os estados mentais inconscientes, a questão que devemos esclarecer é: como exatamente Searle atribui eficácia causal aos estados mentais inconscientes?

Segundo Prata, a resposta pode ser vislumbrada nas passagens em que Searle fala sobre os *fenômenos sensoriais* e os *fenômenos mentais intencionais* (seção 3.5). Mais especificamente, quando o autor ilustra um caso de fenômeno sensorial por meio dos casos de doenças crônicas, como dores nos membros e em partes específicas do corpo. Como vimos, o exemplo de Searle determina que as dores de um indivíduo podem se intensificar durante o sono e com isso despertá-lo no decorrer da noite. A questão que se coloca então é: se os fenômenos mentais

---

<sup>25</sup> Professor do departamento de Filosofia da Universidade Federal de Pernambuco e do Programa de Pós-graduação em Filosofia da mesma instituição.

conscientes são os únicos que possuem poderes causais, como é possível explicar o que acontece no exemplo anterior? Searle esclarece esse ponto através da seguinte passagem:

Qual é exatamente a ontologia da dor quando ela é completamente inconsciente? A resposta me parece óbvia. O que nos predispõe a dizer que a dor continuou a existir, ainda que inconsciente, é que havia um processo neurofisiológico subjacente capaz de gerar um estado consciente e capaz de gerar comportamento apropriado a alguém que tivesse esse estado consciente. (SEARLE, 2006, p. 237)

Essa explicação, no entanto, não satisfaz Prata que em resposta, diz que:

No meu entendimento, isso significa que o processo cerebral, que tem um modo de existência objetivo, isto é, independente das vivências de indivíduos conscientes (cf. SEARLE, 2002, p. 11), está sendo apresentado como uma entidade que tem diversas capacidades, algumas das quais são exercidas, enquanto outras não são. Isso é bastante razoável, pois todo concreto particular possui uma infinidade de propriedades (tanto propriedades atualizadas quanto propriedades disposicionais). A porção de água no interior do meu copo poderia ser usada para dissolver uma certa quantidade de sal, mas ela será usada para saciar a minha sede (temos aqui duas disposições distintas – entre inúmeras outras disposições dessa porção de água, onde apenas uma delas será atualizada). O problema é que se a disposição para gerar um fenômeno consciente é tão distinta, e independente, da disposição de gerar certos comportamentos, isso significa que a consciência não é assim tão decisiva para a existência do fenômeno mental. (PRATA, 2017, p.393)

Com isso, o autor insere uma questão fundamental: se capacidades causais podem ser exercidas por processos neurofisiológicos durante o sono, isso significaria dizer que as propriedades inconscientes não podem ser simplesmente *disposicionais*, afinal, elas *estão causando* um comportamento físico sem que o sujeito esteja consciente. “Se um comportamento é efetivamente causado, em uma situação que nos leva a concluir a presença de uma sensação inconsciente, então temos de concluir que tal sensação existe objetivamente, isto é, independentemente das vivências do sujeito consciente” (PRATA, 2022, p.75). Além disso, se admitirmos que supostamente esses estados possuem eficácia causal, teríamos de admitir também que eles preservam de alguma maneira sua forma aspectual, isto é, as condições de satisfação dos conteúdos intencionais para a efetivação dos estados mentais. Entretanto, isso gera um problema, pois Searle defende que a intencionalidade não pode ser separada da consciência.

Além dos estados sensoriais, o autor explica que é possível encontrar o mesmo



problema nas explicações searliananas sobre os fenômenos intencionais inconscientes. Devemos lembrar que, diferente dos estados não conscientes como, por exemplo, a mielinização, Searle classifica esses fenômenos em formas de intencionalidade inconscientes que são potencialmente acessíveis à consciência, de modo que, se satisfeitos os critérios da forma aspectual, eles podem se tornar estados conscientes.

O ponto chave, assim como no caso dos estados sensoriais, é que os estados intencionais inconscientes dependem de sua relação com a consciência para causar algo. No entanto, se voltarmos ao exemplo das dores, Prata diz que é possível analisá-lo da seguinte maneira: a dor nas costas (um estado sensorial inconsciente) causa o estado intencional de dor; se o indivíduo estiver acordado, a dor é consciente, entretanto, se ele estiver dormindo (e, portanto, inconsciente, já que o sono é um estado de completa inconsciência, cf. SEARLE, 2006) a dor se tornará inconsciente. Sendo assim, podemos concluir que o estado de dor inconsciente é a causa do comportamento que faz o indivíduo acordar. Sobre isso, Prata explica que:

No meu modo de entender, esse problema se coloca de modo especialmente agudo no caso das sensações inconscientes, pois sem consciência não existem qualia, que são aquilo que constitui a sensação do ponto de vista subjetivo, e se o comportamento correspondente à dor continua ocorrendo mesmo na ausência dos qualia, então a causa real do comportamento tem que ser o processo cerebral (objetivo) e não o aspecto qualitativo da sensação (subjetivo). (PRATA, 2017, p.395)

Dessa forma, se concluirmos que esses estados inconscientes não precisam se tornar conscientes para produzir efeitos causais também teríamos de admitir que eles preservam de alguma maneira sua forma aspectual (intencionalidade intrínseca). Mas se este for o caso, a integração entre a intencionalidade e a consciência se tornaria contingente, afinal, “[...] se um fenômeno exerce poderes causais (o que é uma marca da existência) sem estar integrado a esse todo, isso significa que a integração a ele é uma característica secundária, não essencial, daquele fenômeno” (PRATA, 2017, p. 399). Sendo assim, embora a relação entre intencionalidade e consciência faça sentido dentro da argumentação de Searle (haja vista a proposta de centralidade da consciência), o autor não determina satisfatoriamente o modelo de existência dos estados mentais inconscientes, gerando um problema com sua definição de estados mentais inconscientes como disposições.

## 5 CONSIDERAÇÕES FINAIS

Embora tenhamos apresentado três questões fundamentais, a saber, o problema da consciência, a teoria searlina da consciência e a crítica de Searle ao Cognitivismo, em momentos distintos, entendemos que essas questões estão profundamente ligadas ao entendimento do autor de que a consciência é a chave para a resolução dos problemas da mente. De fato, o termo “consciência” é acompanhado de uma enorme ambição em *A Redescoberta da Mente*. Com ele, Searle não apenas propõe uma solução ao problema mente- corpo como também o torna uma medida capaz de apontar quais teorias do mental são adequadas ou não. Assim, a partir das discussões apresentadas até aqui, vimos que Searle mantém as seguintes opiniões:

- (1) Tanto o dualismo quanto o materialismo, e as teorias que decorrem deles têm problemas. Mas é possível traçar um caminho alternativo que leve a uma solução, sobretudo, ao problema mente-corpo;
- (2) A consciência é uma característica biológica de nível superior que atua de forma integrada com a intencionalidade. Em razão disso, todos os estados mentais são definidos em termos dessas duas características;
- (3) O Cognitivismo está comprometido com a ideia de que os estados mentais são estados computacionais intrinsecamente inconscientes;
- (4) A tarefa da ciência não é buscar a objetividade ontológica; sua tarefa é, na verdade, a busca pela objetividade epistemológica, ou seja, a produção de conhecimentos objetivos acerca do mundo.

Para tornar possível (1) Searle propõe (2) sob a alegação (4). Paralelamente, Searle acredita que (3) é consequência de (1) e crítica (3) em favor de (2).

Não parece haver grandes problemas nas alegações de Searle de que existe um problema da consciência, principalmente, em virtude do fracasso materialista em estabelecer uma relação de identidade entre estados conscientes e estados neuronais ou qualquer outro tipo de redução

ontológica. Também consideramos que o QC é uma refutação definitiva da inteligência artificial forte.

No tocante às objeções searlianais ao cognitivismo, podemos dizer que elas são igualmente interessantes, embora a explicação de Castañon seja pertinente. Afinal, se Searle está abertamente criticando o cognitivismo, a noção de cognitivismo que ele utiliza deveria ser a mesma noção usada na ciência cognitiva. Mas, se isso é o caso ou não, a decisão do leitor seria mais justa. E, mesmo admitindo que Searle está errado em suas concepções sobre os cognitivistas, podemos nos perguntar se o erro conceitual de Searle é suficiente para inviabilizar sua crítica ao cognitivismo. De fato, preferimos dizer que, no que concerne à tarefa de responder à questão “o cérebro é um computador digital?”, Searle dá uma resposta satisfatória. Damos especial destaque para a falácia do homúnculo, pois acreditamos que independentemente do alvo de Searle ser o Computacionalismo ou o Cognitivismo, ele traça um limite latente aos computadores que é: *o significado e o modo como um computador processa e transmite informação é puramente resultado da programação humana*. Nós, como sugere a falácia do homúnculo, somos o “agente” “dentro” dos computadores, e por isso o entendimento ou a intencionalidade de sua comunicação é limitado (em termos searlianais) a uma intencionalidade “como-se”. E esse tipo de intencionalidade definitivamente não é suficiente para caracterizar a intencionalidade da mente humana. Em razão disso, concordamos com Searle quanto: não há nenhum homúnculo em nossas cabeças, pois nossos cérebros não são programados para “compreender”, ao contrário, eles compreendem porque são biologicamente constituídos para tal.

Dito isso, a objeção de que existe uma falha nas explicações de Searle sobre a causalidade dos estados mentais inconscientes e sobre o princípio da conexão nos parece justificada. O que decorre da explicação searliana sobre um estado de dor inconsciente é que ele pode ter eficácia causal mesmo quando não é sentido conscientemente como algo “doloroso”. A análise de Prata deixa isso claro: a dor ocorre durante o sono, isto é, em um estado de completa inconsciência. O fato de que o indivíduo acordou incomodado pela dor inconsciente, mostra que um estado objetivo (neurobiológico) que não é consciente nem intencional causou uma ação no mundo, o que se contrapõe conceitualmente ao princípio da conexão.

Admitindo, pois, que há esse problema nas explicações de Searle, quais seriam as implicações disso para nossa discussão? A implicação mais óbvia é que a solução de Searle para

o problema da consciência se torna fragilizada. Essa solução tem um custo alto, afinal, ela depende fundamentalmente da sustentação da centralidade da consciência.

No que se refere a suas alegações contra o cognitivismo, acreditamos que a dificuldade reside na forma como Searle defende enfaticamente que não existem estados mentais puramente inconscientes, ou seja, estados mentais (sejam eles computacionais ou não) que nunca aparecem sob a luz da consciência. E essa dificuldade implica em outra quando Searle apela ao princípio da conexão para afirmar que somente estados mentais conscientes têm poder causal, gerando fenômenos ou ações no mundo. Dessa forma, ainda que a explicação computacional da mente incorra de fato na falácia do homúnculo, a sugestão cognitivista de que os estados mentais são causados por regras internas não seria refutada, ao menos não pelo princípio da conexão.

## REFERÊNCIAS

- BARROSO, C. A. C. As Ciências Cognitivas e a Metáfora do Computador. In: Marcus José Alves de Souza e Maxwell Morais de Lima Filho. (Org.). **Escritos de Filosofia II - Linguagem e Cognição**. 1ed.Maceió: EDUFAL, 2017, v, p. 317-338.
- BENNET, M. R; HACKER, P. M. S. **Fundamentos filosóficos da neurociência**. Lisboa: Instituto Piaget, 2005.
- CARNAP, R; HAHN, H; NEURATH, O: A concepção científica do mundo: o círculo de Viena. In: **Cadernos de História e Filosofia da Ciência**. v.10, p. 5-20, 1986. Disponível em: <https://www.cle.unicamp.br/eprints/index.php/cadernos/issue/view/216> Acesso em: 8 fev. 2022.
- CASTAÑON, G. A. **O Cognitivismo e o desafio da Psicologia Científica**. Rio de Janeiro: UFRJ/IP, 2006a.
- CASTAÑON, G. A. **John Searle e o cognitivismo**. Ciênc. cogn., Rio de Janeiro, v. 8, p. 96-109, ago, 2006b.
- CHALMERS, A, F. **O que é a ciência afinal?** São Paulo: Editora Brasiliense,1993.
- CHURCHLAND, P, M. **Matéria e Consciência: uma introdução contemporânea à filosofia da mente**. Tradução: Maria Clara Cescato. São Paulo: Editora UNESP, 2004.
- COSTA, C. **Filosofia da mente**. Rio de Janeiro: Jorge Zahar Editor, 2005.
- DREYFUS, H. L. **O que os computadores não podem fazer: uma crítica da razão artificial**. Rio de Janeiro. A Casa do Livro Eldorado, 1975.
- FODOR, J, A. **O problema mente-corpo**. TCFC3 – Filosofia das Ciências Neurais, São Paulo: FFLCH-USP, 2011.
- GARDNER, H. **A nova ciência da mente: uma história da revolução cognitiva**. São Paulo: EDUSP, 3. ed, 2003.
- HODGES, A. **Turing: um filósofo da natureza**. São Paulo: UNESP, 2001.
- JAPIASSU, H. MARCONDES, D. **Dicionário Básico de Filosofia**. Rio de Janeiro: Jorge Zahar Editor, 2001.
- LECLERC, A. **Uma introdução à filosofia da mente**. Curitiba: Appris, 2018.
- LIMA, S. N. S. **Searle e a inteligência artificial: uma crítica a partir do Naturalismo Biológico**. Monografia (Graduação) – Universidade Estadual do Piauí - UESPI, Curso

Licenciatura Plena em Filosofia, Campus Prof. Alexandre Alves de Oliveira, Parnaíba-PI, 2020.

LUGER, G. F. **Inteligência artificial: estruturas e estratégias para a solução de problemas complexos**. Tradução Paulo Martins Engel. Porto Alegre: Bookman, 2007.

MIGUENS, S. **Compreender a Mente e o Conhecimento**. Lisboa: Porto, 2009.

NAGEL, T. **Como é ser um morcego?** Tradução: Josemar de Campos Maciel. Revista da Abordagem Gestáltica - Phenomenological Studies – XIX (1): 109-115, jan-jul, 2013.

OLIVEIRA, M. B. **Cognitivism and cognitive science**. Trans/Form/Ação, São Paulo, 13: 85-93, 1990.

PESSOA JR., O. **Eventos quânticos e reducionismo causal**. Principia, v. 17, p. 365-381, 2013.

REALE, G. **História da filosofia: de Freud à atualidade**. São Paulo: Paulus, 2006.

PRATA, T. A. **A teoria disposicional de Searle e o problema da causação mental inconsciente**. Revista Filosófica de Coimbra, vol. 31, nº61, 2022.

PRATA, T. A. “Uma crítica à concepção disposicional de Searle sobre os fenômenos mentais inconscientes”. In: Araújo, A.; Correia, A.; Ghiraldelli Jr., P.; Leclerc, A.; Mograbi, G. **Pragmatismo, filosofia da mente e filosofia da neurociência**. Coleção XVII Encontro ANPOF: ANPOF, p. 387-404, 2017.

PRATA, T. A. **Um argumento contra a tese da subjetividade ontológica da consciência no naturalismo biológico de John Searle**. Unisinos Journal of Philosophy 21(3):303-311, sep/dec 2020.

RYLE, G. “O mito de Descartes”. In: M. Luísa Nunes. **Introdução à psicologia: o conceito de espírito**. Lisboa: Moraes Editora, 1970, p. 13-25. Disponível em: [www.fflch.usp.br/df/opessoa/Ryle-Mito-Descartes-2.pdf](http://www.fflch.usp.br/df/opessoa/Ryle-Mito-Descartes-2.pdf). Acesso em: 20 fev. 2022.

SCHWARTZ, S. P. **Uma breve história da filosofia analítica: de Russell a Rawls**. São Paulo: Loyola, 2017.

SEARLE, J. R. **A redescoberta da mente**. São Paulo: Martins Fontes, 2006.

SEARLE, J. R. **Consciousness, Explanatory Inversion, and Cognitive Science**. Behavioral and Brain Sciences, 13(1): 585-642, 1990

SEARLE, J. R. **Intencionalidade**. São Paulo: Martins fontes, 2002.

SEARLE, J. R. **Mente, cérebro e ciência**. Lisboa: Edições 70, 2017.

- SEARLE, J. R. **Mente, linguagem e sociedade**: filosofia no mundo real. Rio de Janeiro:Rocco, 2000.
- SEARLE, J. R. **Minds, brains, and programs**. In Behavioral and Brain Sciences 3 (3): 417-457, 1980.
- SKINNER, B.F. **O comportamento verbal**; tradução de Maria da Penha Villalobos. — São Paulo: Cultrix: Ed. da Universidade de São Paulo, 1978.
- STERNBERG, R. J. **Psicologia cognitiva**. Porto Alegre: Artmed, 2008.
- TEIXEIRA, J. F. **Mente, cérebro e cognição**. Petrópolis, RJ: Vozes, 2011.
- TEIXEIRA, J. F. **Mentes e máquinas: uma introdução à ciência cognitiva**. Porto Alegre: Artes Médicas, 1998.
- TURING, A.M. “Computadores e inteligência”. In: **Cibernética e comunicação**. São Paulo: Editora Cultrix, 1973.