



# O QUE É FILOSOFIA DA MENTE

João de Fernandes Teixeira

2ª Edição



A filosofia da mente ainda é um território desafiador para os filósofos profissionais. Ela não é uma reflexão feita no conforto do a priori ou do comentário aos textos de filósofos consagrados pela história. Na agenda dos filósofos da mente há temas derivados de ciências empíricas como a ciência da computação, a linguística e a neurociência. A filosofia da mente herda do método científico a revisão constante dos resultados das pesquisas e, juntamente com ela, a necessidade de readaptar seus temas de discussão e suas prioridades.

**João de Fernandes Teixeira** é um dos pioneiros da filosofia da mente no Brasil. Bacharel em filosofia pela USP e mestre em filosofia da ciência pela UNICAMP é também PhD pela University of Essex, na Inglaterra. Fez pós-doutorado nos Estados Unidos, sob orientação de Daniel Dennett. Foi colaborador do Instituto de Estudos Avançados da USP e lecionou em várias universidades brasileiras como a UNESP, a UFSCar e a PUC-SP. Publicou 14 livros na área de filosofia da mente e ciência cognitiva. Mantém a página Filosofia da Mente no Brasil no Facebook.



# **O QUE É FILOSOFIA DA MENTE**



# Filosofia & Interdisciplinaridade

- **Agnaldo Cuoco Portugal**, UNB, Brasil
- **Alexandre Franco Sá**, Universidade de Coimbra, Portugal
- **Christian Iber**, Alemanha
- **Claudio Gonçalves de Almeida**, PUCRS, Brasil
- **Cleide Calgato**, UCS, Brasil
- **Danilo Marcondes Souza Filho**, PUCRJ, Brasil
- **Danilo Vaz C. R. M. Costa**, UNICAP/PE, Brasil
- **Delamar José Volpato Dutra**, UFSC, Brasil
- **Draíton Gonzaga de Souza**, PUCRS, Brasil
- **Eduardo Luft**, PUCRS, Brasil
- **Ernildo Jacob Stein**, PUCRS, Brasil
- **Felipe de Matos Muller**, UFSC, Brasil
- **Jean-François Kervégan**, Université Paris I, França
- **João F. Hobuss**, UFPEL, Brasil
- **José Pinheiro Pertille**, UFRGS, Brasil
- **Karl Heinz Efken**, UNICAP/PE, Brasil
- **Konrad Utz**, UFC, Brasil
- **Lauro Valentim Stoll Nardi**, UFRGS, Brasil
- **Marcia Andrea Bühring**, PUCRS, Brasil
- **Michael Quante**, Westfälische Wilhelms-Universität, Alemanha
- **Miguel Giusti**, PUCP, Peru
- **Norman Roland Madarasz**, PUCRS, Brasil
- **Nythamar H. F. de Oliveira Jr.**, PUCRS, Brasil
- **Reynner Franco**, Universidade de Salamanca, Espanha
- **Ricardo Timm de Souza**, PUCRS, Brasil
- **Robert Brandom**, University of Pittsburgh, EUA
- **Roberto Hofmeister Pich**, PUCRS, Brasil
- **Tarcílio Ciotta**, UNIOESTE, Brasil
- **Thadeu Weber**, PUCRS, Brasil

# O QUE É FILOSOFIA DA MENTE

2ª. Edição

João de Fernandes Teixeira



**Direção editorial:** Agemir Bavaresco  
**Diagramação e capa:** Lucas Fontella Margoni  
**Arte de capa:** Romain Briaux

**O padrão ortográfico e o sistema de citações e referências bibliográficas são prerrogativas de cada autor. Da mesma forma, o conteúdo de cada capítulo é de inteira e exclusiva responsabilidade de seu respectivo autor.**



Todos os livros publicados pela Editora Fi estão sob os direitos da [Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/deed.pt_BR)  
[https://creativecommons.org/licenses/by/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by/4.0/deed.pt_BR)



Associação Brasileira de Editores Científicos

<http://www.abecbrasil.org.br>

Série Filosofia e Interdisciplinaridade — 58

Dados Internacionais de Catalogação na Publicação (CIP)

TEIXEIRA, João de Fernandes.

O que é filosofia da mente. [recurso eletrônico] / João de Fernandes Teixeira. 2a ed. -- Porto Alegre, RS: Editora Fi, 2016.

47 p.

ISBN - 978-85-5696-066-5

Disponível em: <http://www.editorafi.org>

1. Filosofia da mente. 2. Epistemologia. 3. Filosofia da ciência. 4. Consciência. 5. Problema corpo e mente. I. Título. II. Série.

CDD: 150

Índices para catálogo sistemático:

1. Filosofia da mente 150

**Para Malu, Jujuba e Tâmara.**





*O filósofo é uma pessoa que chuta a terra com força,  
levanta muita poeira e depois reclama  
porque não consegue enxergar.*

Leibniz



# PREFÁCIO A SEGUNDA EDIÇÃO

---

Publiquei este livro pela primeira vez em 1994, na coleção Primeiros Passos, da Editora Brasiliense. Naquela época pouco se falava de filosofia da mente no Brasil.

No ano anterior eu tinha aberto uma linha de pesquisa em filosofia da mente e ciência cognitiva na Universidade Federal de São Carlos, na pós-graduação em Filosofia. Era uma linha de pesquisa pioneira no Brasil, para a qual afluíram estudantes de diversas áreas do conhecimento: a matemática, a psicologia, a ciência da computação e a psiquiatria. Mas os estudantes formados em filosofia mantiveram sua habitual distância prudente em relação às novidades.

A filosofia da mente ainda é um território desafiador para os filósofos profissionais. Ela não é uma reflexão feita no conforto do *a priori* ou do comentário aos textos de filósofos consagrados pela história. Na agenda dos filósofos da mente há temas derivados de ciências empíricas como a ciência da computação, a linguística e a neurociência. A filosofia da mente herda do método científico a revisão constante dos resultados das pesquisas e, juntamente com ela, a necessidade de readaptar seus temas de discussão e suas prioridades.

Existe, até hoje, muito preconceito em relação à filosofia da mente. Preconceito significa tomar uma posição sem conhecer um determinado assunto. Ou seja, o que não se conhece se destrói ou, simplesmente, se desqualifica.

Espero, com este pequeno livro, diminuir esse preconceito e mostrar que, apesar de suas peculiaridades, a filosofia da mente está profundamente mergulhada nos temas filosóficos tradicionais.



# SUMÁRIO

---

A NATUREZA DO MENTAL

15

O PROBLEMA MENTE-CÉREBRO

20

MAQUINISMO E CONSCIÊNCIA

28

A IDENTIDADE PESSOAL

37

CONCLUSÃO

45



# A NATUREZA DO MENTAL

---

Durante séculos os filósofos tentaram responder às questões: O que é a mente? O que caracteriza os fenômenos mentais? O mesmo ocorre com quase todas as religiões que conhecemos. Todas elas referem-se à mente às vezes como "espírito" ou como "alma" - algo que teria propriedades especiais e que continuaria subsistindo mesmo após a nossa morte.

Na verdade, falar de "mente" ou de "fenômenos mentais" ainda é uma coisa que nos causa tanta estranheza quanto falar de OVNI's ou da existência de criaturas extraterrestres. A mente sempre foi um enigma, talvez pelo fato de os fenômenos mentais serem invisíveis e inacessíveis para nós. A ciência de que dispomos até hoje não parece ter auxiliado muito na tentativa de encontrar uma resposta para essas questões. A Psicologia quer fazer uma ciência da mente, desenvolveu testes e teorias acerca do funcionamento mental do homem e de alguns animais. Mas os psicólogos nunca chegaram a um consenso sobre *o que é a mente* e sobre *o que* eles estão falando. Há não muito tempo havia psicólogos que nem sequer reconheciam a existência da mente ou dos fenômenos mentais, embora se declarassem estudiosos de Psicologia. Estranhas criaturas, que nem sequer sabiam que sonhavam!

Por outro lado, nos últimos anos presenciamos um imenso desenvolvimento da Neurofisiologia e das chamadas "ciências de cérebro". Mas será que essas ciências podem nos ajudar a encontrar uma resposta para tais questões? Podemos até imaginar um neurocirurgião abrindo a cabeça de alguém e examinando seu cérebro: certamente ele verá muitas células nervosas, mas nunca verá uma ideia, um sentimento ou uma emoção. Talvez esta seja uma maneira de caracterizar a natureza dos fenômenos mentais: eles são invisíveis. Mas será esta uma boa caracterização?

Os átomos também são invisíveis. Entretanto seria difícil dizer que átomos são fenômenos mentais apenas porque não podemos observá-los. A diferença estaria no fato de que os átomos são invisíveis, mas isto não significa que um dia eles não *possam ser observados*. Através de um supermicroscópio eletrônico, que poderia ser construído no futuro - quem sabe? O mesmo não ocorreria com os fenômenos mentais. Assim, parece absurdo supor que um dia, mesmo com um supermicroscópio eletrônico, poderemos observar uma ideia, um sentimento ou uma emoção.

A esta altura o neurobiólogo ou estudioso das "ciências do cérebro" poderia objetar: "É claro que eu só observo células nervosas quando abro a cabeça de alguém. Não posso observar ideias, sentimentos ou emoções da mesma maneira que o físico não pode observar a "massa", "aceleração" ou "gravidade". Isto não quer dizer que o estudo da mente não possa ser feito a partir do estudo do cérebro. Mas ao fazer esta afirmação o neurobiólogo estaria se esquecendo de uma diferença fundamental: "massa", "aceleração", "gravidade" podem ser medidas. Não teria sentido dizer que um dia poderemos "medir uma ideia" ou estabelecer a "quantidade de alegria" que sentimos ao descobrir que nosso bilhete de loteria foi premiado. Falamos de coisas que nos dão mais alegria ou menos alegria, mas isto está longe de ser uma medida das emoções. Seria estranho pensar que um dia poderíamos ter uma escala para medir alegrias: hoje minha alegria está no nível 1,8; depois de alguns copos de vinho ela poderá chegar no nível 3,6. A alegria que sinto quando encontro o bilhete de loteria premiado não é maior nem menor do que aquela que sinto quando a menina que eu paquerava há meses me dá um beijo: elas são alegrias *diferentes*.

Fenômenos mentais não são apenas invisíveis e impossíveis de serem medidos. A grande diferença estaria no fato de eles serem inacessíveis à observação. Se digo que



tenho febre, alguém pode colocar um termômetro sob o meu braço e, após alguns minutos, afirmar: "Não, você não tem febre, o termômetro marca apenas 36,5 graus". Mas se digo "Estou com dor de cabeça", ninguém, exceto eu, pode saber se estou mentindo ou dizendo a verdade. Alguém pode dizer e fingir que está com dor de cabeça: somente essa pessoa saberá se o que diz é verdade ou não. A inacessibilidade dos fenômenos mentais torna-os essencialmente *subjetivos* ou *privados*, para usar o jargão filosófico.

Ora, se não podemos observar uma ideia, um sentimento ou uma emoção, porque são imateriais, isto equivale a dizer que pensamentos não podem ser destruídos: só podemos destruir coisas que estão no mundo. Seria possível aniquilar criaturas e seus cérebros, mas não ideias e pensamentos. Dessa suposição originou-se a concepção de que a mente (ou a alma) seria imortal e que persistiria no tempo, mesmo após a desintegração do corpo. Tudo se passa como se alguém dissesse: é possível destruir tudo o que é triangular, mas você nunca poderá destruir a ideia de triângulo. Foi com base num argumento parecido com este que alguns filósofos sustentaram, durante muito tempo, a imortalidade da mente (ou da alma).

Mas serão a persistência no tempo e a indestrutibilidade características exclusivas dos fenômenos mentais? Será que alguém, alguma vez, conseguiu destruir um pedaço de matéria? "Nada se cria, tudo se transforma". Nem mesmo uma bomba atômica pode destruir um pedaço de matéria; ela pode, no máximo, *transformá-la*. A indestrutibilidade não seria, assim, uma característica exclusiva da mente.

Fenômenos mentais são subjetivo e privados, o que significa dizer que eles ocorrem para nós. Mas será isto suficiente para manter a afirmação de que a mente nada tem a ver com o cérebro? Tenho uma dor de cabeça e

tomo uma aspirina; se a dor passa, em que local terá atuado a aspirina? Na mente ou no cérebro? A dor pode ser algo privado e inacessível, mas como pode uma substância química influir na mente? Não parece claro que indivíduos com grandes porções de seus cérebros danificados em acidentes perdem igualmente grande parte de suas atividades mentais? Mente e cérebro podem ser coisas distintas, mas certamente estão ligados de alguma maneira. O problema é saber como é possível essa ligação - e este é um problema para o filósofo da mente resolver.

Os filósofos da mente sempre tiveram por objetivo esclarecer questões fundamentais, tais como: O que distingue a mente de outros objetos que estão no universo? Qual a natureza do pensamento? Será o pensamento algo imortal e eterno? Serão mente e cérebro uma só e mesma coisa? Será a distinção entre espírito e matéria apenas uma ilusão produzida pela nossa linguagem ou pela nossa cultura?

Na verdade, essas sempre foram questões da Filosofia e percorrem toda sua história. Mas de 1940 para cá elas começaram a ganhar uma ênfase especial por parte de alguns filósofos, principalmente na Inglaterra e nos Estados Unidos. A década de 1940 presenciou grandes inovações científicas e tecnológicas. Os estudos sobre o cérebro tinham avançado muito e já surgiam os primeiros computadores digitais. Falava-se de uma analogia entre computadores e cérebro, e isto forçava os filósofos a refletirem sobre uma afirmação que se tornava cada vez mais frequente entre cientistas: a de que uma máquina pode pensar. É claro que tal afirmação tinha ecos mais profundo entre os filósofos. Tratava-se de saber, afinal de contas, o que é essa coisa estranha que chamamos de pensamento.

Os progressos da neurofisiologia também atormentavam os filósofos. Experimentos com o cérebro, avanços em técnicas neurocirúrgicas abriram uma perspectiva fantástica: a de que um dia seria possível

realizar um transplante de cérebro. Se isto fosse possível, diziam alguns, todas as questões da Filosofia da Mente estariam resolvidas. Seria possível verificar se o organismo transplantado adquiria as mesmas memórias e pensamentos do doador. Neste caso, a mente não seria nada mais do que o cérebro.

O transplante de cérebro nunca foi realizado e é provável que nunca seja possível. Talvez não seja esse o caminho para sabermos alguma coisa acerca da natureza da mente. Mas a possibilidade deste tipo de cirurgia reavivava uma das questões mais interessantes da Filosofia da Mente: o que permite dizer que eu sou eu, ou, em outras palavras, o que me confere identidade pessoal? Serão minhas memórias? Será a estrutura química peculiar de meu cérebro?

A Filosofia, no início do século XX, tinha seguido uma trajetória bastante estranha. A grande preocupação com a linguagem, que era então dominante, sugeria que todos os problemas da Filosofia da Mente nada mais eram do que ilusões. Ilusões produzidas pela própria linguagem, que teria se tornado um imenso labirinto no qual a reflexão teria se enredado e se perdido. Seria inútil refletir sobre a mente sem antes refletir sobre a linguagem. Mas esta tendência foi sendo progressivamente revertida. Hoje em dia os filósofos da mente não desprezam a reflexão sobre linguagem, mas estão convencidos de que ela, por sim só, é apenas um ponto de partida para iniciar a discussão dos problemas que estão envolvidos na descoberta da natureza e das propriedades dos fenômenos mentais.

# O PROBLEMA MENTE-CÉREBRO

---

Foi Descartes (1596-1650) que, pela primeira vez, formulou explicitamente a necessidade de se distinguir entre mente e corpo. Claro que outros filósofos, desde a antiguidade, já haviam refletido sobre a natureza da alma (ou da mente) e apontado para aquilo que julgavam ser algumas de suas características especiais, como, por exemplo, a imaterialidade e a imortalidade. Mas a filosofia de Descartes (o cartesianismo) serviu para reacender um debate que atravessa todo o pensamento moderno: a polêmica entre o *monismo* e o *dualismo*.

O *monismo* é a tese que sustenta que só existe um tipo de substância no universo, seja ela material ou espiritual. A versão mais frequente do monismo é o *materialismo*, ou seja, a teoria de que não existe nada além da matéria e suas possíveis manifestações no universo. De acordo com essa visão, fenômenos mentais são idênticos aos fenômenos físicos, pois mente e cérebro são a mesma coisa. Por outro lado, o *dualismo* sustenta que há duas substâncias do universo e que existe uma diferença fundamental e irreconciliável entre elas. Nunca poderíamos supor que a mente e cérebro são a mesma coisa. A versão mais conhecida do dualismo é o que chamamos de *espiritualismo*.

Descartes era um dualista. Ele supunha que mente e matéria tinham propriedades radicalmente diferentes. Um pedaço de matéria, por menor que fosse, seria sempre divisível. O mesmo não poderíamos afirmar acerca de uma ideia ou de um estado mental: não teria cabimento supor que um dia poderíamos dividir um pensamento em fatias, da mesma forma que fazemos com um pedaço de pão ou uma barra de ferro. Mesmo quando temos uma ideia complexa e procuramos transformá-la em várias ideias

simples, cada uma delas será sempre uma unidade indivisível.

Esse argumento era bastante convincente e exerceu grande fascínio entre os filósofos contemporâneos de Descartes. Afinal, encontrava-se uma boa razão para postular a existência de uma assimetria entre mente e corpo. Mas logo algumas dificuldades começaram a ser apontadas. Se alma e corpo são distintos e se a mente é imaterial, como poderia ela influir sobre nossas ações? Nunca poderíamos imaginar que algo imaterial pode afetar alguma coisa do mundo como, por exemplo, nosso próprio corpo. Mas isto certamente vai contra nossas intuições mais simples e cotidianas: temos certeza de que é a nossa mente, com seus pensamentos, desejos e intenções, que regula e causa nossos comportamentos ou ações. Separar mente e corpo parecia uma tarefa relativamente fácil: os problemas começavam a aparecer na medida em que os dualistas precisavam encontrar uma maneira de conceber o modo como essas duas coisas poderiam interagir. Se a mente é separada do corpo, como posso saber qual é o *meu* corpo?

Descartes tentou oferecer uma solução para esses problemas. Ele falava da existência de um órgão especial, localizado no cérebro, a glândula pineal. Este órgão (que hoje sabemos ser a hipófise) seria responsável por estabelecer uma ponte entre a alma e o corpo. Mas como isso seria possível, Descartes nunca explicou. Para estabelecer essa ligação a glândula pineal teria de estar a meio caminho entre algo material e algo imaterial. Seria possível conceber a existência de um órgão com tais características? A filosofia de Descartes acabava se tornando um manual de anatomia fantástica!

A partir desse episódio apareceram várias teorias tentando relacionar mente e corpo. A mais interessante e original talvez tenha sido a do filósofo alemão Gottfried Wilhelm Leibniz. Ele acreditava na existência de uma harmonia preestabelecida no universo. Mente e corpo não

precisariam ter nenhum tipo de ligação, pois, de acordo com a harmonia preestabelecida, tudo o que se passa na esfera do mental encontra um correspondente na esfera do mundo físico. O físico e o mental não precisam ter nenhuma ligação entre si, eles apenas "caminham juntos" como se no início do universo um Deus tivesse programado o mundo ao modo de duas séries que correm simultânea e harmoniosamente. Mas essa teoria soa hoje, para nós, no mínimo como algo bizarro.

Por não conseguir boas soluções teóricas para o problema da ligação entre mente e corpo, as teorias filosóficas oscilaram, a partir de então, entre posições radicalmente monistas (materialistas) ou posições radicalmente dualistas. Mas sempre havia dificuldades nas tentativas de reduzir o mental ao físico ou para sustentar que a mente nada tem a ver com o cérebro.

No século XIX esse debate parece ter perdido um pouco o fôlego inicial. A filosofia se ocupava de outras questões que pareciam mais urgentes, como, por exemplo, os destinos da humanidade, da história e dos conflitos sociais. Ademais, a influência da filosofia de Immanuel Kant (1724-1804) ainda era muito grande. Kant julgava ter mostrado, de uma vez por todas, que nunca poderíamos chegar a uma solução para o problema das relações mente-corpo nem tampouco chegar a uma conclusão definitiva acerca da natureza do pensamento. Essas questões seriam, rigorosamente falando, indecidíveis, e seria inútil tentar defender seja o dualismo seja o monismo, uma vez que sempre haveria razões igualmente fortes para adotar uma posição ou outra.

Foi preciso aguardar os progressos da Neurofisiologia para que este debate se reacendesse. O final do século XIX foi particularmente fértil para essas investigações. Descobriu-se o neurônio e sua capacidade de transmitir energia elétrica. As pesquisas sobre anatomia do cérebro também viveram progressos consideráveis. E no

final do século XIX iniciam-se pesquisas mais sistemáticas sobre a natureza das doenças mentais e do hipnotismo. Surge a psicanálise de Sigmund Freud, e com ela os médicos e cientistas voltam a se perguntar se mente e cérebro seriam uma só e mesma coisa. As investigações de Freud e sua aplicação na tentativa de curar alguns transtornos mentais abriram perspectivas novas e desconhecidas acerca da natureza da mente.

Mas é no século XX que vai surgir a Filosofia da Mente propriamente dita. A Filosofia da mente é um novo esforço para retornar os principais temas clássicos que atravessam o pensamento na modernidade. Era preciso fazer uma nova tentativa no sentido de determinar a natureza última dos fenômenos mentais; uma tentativa que faria a reflexão filosófica mergulhar novamente em direção ao exame das grandes teorias metafísicas, mas que não poderia, dessa vez, ignorar os resultados das pesquisas sobre o cérebro humano. A questão das relações entre mente e cérebro passa a construir uma de suas preocupações fundamentais. Era preciso encontrar novas teorias que pudessem dar conta das relações entre fenômenos físicos e fenômenos mentais. Esboçar tais teorias era necessidade premente, sobretudo porque o século XX tinha se iniciado com uma forte tendência para a adoção do monismo materialista, resultante do grande entusiasmo pelas pesquisas neurofisiológicas que se avolumavam cada vez mais.

Ao abordar o problema das relações mente-cérebro os filósofos da mente tentaram inovar, propondo novas teorias. Uma das teorias mais notáveis foi proposta por Gilbert Ryle em 1949. Ela marcou o início da Filosofia da Mente contemporânea. Ryle dizia que o problema das relações mente-corpo não deveria sequer ser considerado um autêntico problema; ele seria o resultado de uma imensa confusão teórica cuja origem está na maneira como empregamos nossa linguagem.

Ao longo dos séculos nossa cultura teria gerado, equivocadamente dois tipos de vocabulários: um "vocabulário do físico" e um "vocabulário do mental". Os dois falam exatamente da mesma coisa, mas seu emprego errôneo levou-nos a crer que o físico e o mental são duas substâncias diferentes e incompatíveis. Se procedermos a uma análise rigorosa do emprego da nossa linguagem veremos, no final, que nem faz sentido falar de um problema mente-corpo. Este problema seria uma ilusão que deveria ser dissipada, e esta seria uma tarefa que o filósofo da mente teria de empreender através da análise linguística.

Por exemplo, quando dizemos "minha mente está cansada de tanto estudar" ou "estes pensamentos me causam dor de cabeça" estaríamos alimentando tal ilusão, pois essas são expressões que implicitamente se referem à mente como uma coisa ou uma substância concebível como algo separado do corpo. Nossa linguagem está povoada por centenas de expressões desse tipo: por isso teria surgido o "problema mente-corpo" que, na verdade, nunca teria sido mais do que um grande equívoco dos filósofos anteriores.

Mas o entusiasmo pela teoria de Ryle durou pouco. Uma forte tendência em direção ao materialismo começou a se manifestar nas décadas de 1950 e 1960. Uma verdadeira onda de ensaios e artigos escritos por filósofos americanos e ingleses invadiu as revistas filosóficas da época. Aliás, os filósofos da mente sempre preferiram os ensaios e pequenos artigos sobre temas específicos, evitando, na maioria das vezes, a produção de livros mais extensos.

Novas versões do materialismo foram propostas e passaram a predominar na Filosofia da Mente. Os novos materialistas eram muito imaginosos e logo propuseram uma teoria que ficou conhecida como *teoria da identidade entre mente e cérebro*. Essa teoria foi proposta por um filósofo australiano, J.J.C.Smart, que sustentava que estados mentais



são idênticos a estados cerebrais. Ambos seriam a mesma coisa, com uma diferença apenas aparente, da mesma maneira que as nuvens e gotículas de água, embora seja comum nos referirmos a elas como elementos distintos.

Mas será cabível sustentar uma identidade entre mente e cérebro? Se duas coisas são idênticas, elas devem ter as mesmas propriedades. Estados cerebrais se devem a mudanças que ocorrem com os neurônios. Os neurônios são úmidos, transmissores de corrente elétrica e ocupam uma posição no espaço. Terá sentido supor que estados mentais poderiam ter essas mesmas propriedades? Teria sentido afirmar que meu pensamento é úmido, ou que minha ansiedade está a cinco centímetros do hemisfério direito do meu cérebro?

Os materialistas precisavam imaginar uma outra maneira de conceber a identidade entre estados mentais e estados cerebrais. Uma solução interessante foi o que eles chamaram de *identidade teórica*. Um exemplo de identidade teórica é a seguinte afirmação: água = H<sub>2</sub>O. Hoje em dia aceitamos esse enunciado com naturalidade, mas isso só passou a fazer sentido depois que se fez a análise química da água e se descobriu que ela é composta por duas partes de hidrogênio e uma parte de oxigênio. Da mesma maneira afirmamos que a luz é radiação eletromagnética. Isto passou a fazer sentido depois da descoberta de todo um conjunto de teorias físicas que julgamos serem verdadeiras. Ora, dizem os materialistas, talvez um dia sejam descobertas teorias neurofisiológicas a partir das quais passe a fazer sentido a afirmação - embora atualmente ainda nos pareça estranha - de que estados mentais são estados cerebrais. Mas será que isso resolve todos os nossos problemas? O que nos garante que um dia surgirão tais teorias?

O materialismo é uma doutrina que enfrenta inúmeras dificuldades. O entusiasmo por teorias materialistas surgiu do fato de que cientistas e filósofos um

dia acreditaram que seria possível encontrar um correspondente cerebral para cada um de nossos estados mentais. Os aparelhos de eletroencefalografia, que permitem medir os impulsos elétricos do cérebro e representá-los numa fita de papel, devem ter contribuído fortemente para difundir essa crença. Tornou-se possível, nos modernos laboratórios de pesquisas, saber quando um indivíduo que dorme começa a sonhar, bastando para isso observar os gráficos produzidos pelos aparelhos de eletroencefalografia. Mas, mesmo que esses aparelhos permitam saber quando uma pessoa está sonhando, eles não permitem saber *com o que* essa pessoa está sonhando. Da mesma maneira, se analisarmos as modificações químicas do cérebro de uma pessoa, poderemos saber se essa pessoa está tendo um acesso de raiva, mas essa análise não nos permite saber *do que* ela tem raiva. Há muito mais estados mentais do que seus correspondentes cerebrais, apesar de termos bilhões de neurônios formando inúmeras ligações e combinações entre si. Uma redução de estados mentais a estados cerebrais parece uma tarefa quase impossível, e as tentativas de realizá-la só têm servido para nos afastar do materialismo. Ao contrário, elas parecem mostrar que pensamentos são privados e inacessíveis, ou seja, algo que escapa dos limites do mundo material.

Devemos então abandonar o materialismo? Talvez seja melhor reformular essa questão de um modo ainda mais radical: é possível ser materialista? Se todos os meus estados mentais são resultados das relações químicas do meu cérebro, como quer o materialista, então devo supor que o pensamento "estados mentais= estados cerebrais" *também* é resultado dessas reações químicas. Isto significa que, se nos próximos segundos a base química de meu cérebro mudar, eu poderia passar a sustentar o ponto de vista oposto. O materialismo torna-se, assim, uma tese no mínimo autocontraditória!

Ora, qual será a alternativa ao materialismo? O dualismo? O leitor já deve ter suspeitado de que falamos muito pouco do dualismo. Mas isto não se deve ao fato de querermos ser propositadamente tendenciosos. Se adotarmos o dualismo, praticamente não nos resta nada para falar, exceto a afirmação de que mente e cérebro são coisas distintas. O dualismo não nos diz nada acerca da natureza da mente, apenas afirma que ela não é material. Ele só nos fornece um ponto de partida, e isto é muito pouco para que possamos elaborar uma ciência ou uma filosofia da mente.

O que fazer então? Seria preciso encontrar uma alternativa, seja ao materialismo, seja ao dualismo. Mas isso seria o mesmo que querer encontrar a terceira margem de um rio - algo que talvez a Filosofia da Mente nunca possa fazer.

# MAQUINISMO E CONSCIÊNCIA

---

Algumas décadas atrás, o matemático inglês Alan Turing formulou a seguinte pergunta: Pode uma máquina pensar? Nessa época o computador digital acabava de ser inventado e vivia-se na Europa uma fase de grande otimismo em relação ao que os computadores poderiam fazer. Nascia a Inteligência Artificial, uma disciplina nova que tinha por objetivo o estudo e a construção de máquinas pensantes. Dizia-se que em poucos anos seria possível construir uma máquina que pensa. Isto fez com que os filósofos voltassem a se debruçar sobre algumas questões centrais da Filosofia da Mente: o que é pensamento? Será ele exclusividade dos seres humanos e de alguns animais? O que impede de atribuir a um computador a capacidade de pensar? Caso um computador chegue a pensar, poderá ele adquirir consciência do que está pensando e com isto igualar-se a um ser humano?

Essas questões na verdade não são novas. No século XVII Descartes já se preocupava com esses problemas. Os cientistas da época já falavam da construção de autômatos e se questionavam sobre o que poderia impedir alguém de um dia construir uma máquina pensante. Descartes tinha uma série de restrições a esse projeto. Ele dizia que, por mais que os autômatos fossem aperfeiçoados, eles nunca poderiam ter uma vida mental igual à dos seres humanos.

Descartes era um dualista. Ele acreditava que o pensamento era uma propriedade da alma imaterial, do espírito, e sobre este tema ele tinha escrito uma longa demonstração que começava na dúvida radical e terminava na sua famosa frase "Penso, logo existo". Supor que uma máquina poderia pensar seria o mesmo que abandonar o dualismo. Afinal de contas, isso é o mesmo que supor que algum tipo de arranjo de elementos materiais pode produzir

pensamento sem precisar supor a existência de uma substância mental ou espiritual. Mas era precisamente neste ponto que Descartes identificava uma diferença fundamental e intransponível entre autômatos e seres humanos: mesmo que se construísse um autômato extremamente aperfeiçoado ou uma réplica perfeita que fizesse tudo que nós podemos fazer, a essas criaturas estaria sempre faltando a substância espiritual.

No século XX esta discussão é retomada por Alan Turing. Ela adquire grande importância neste contexto, uma vez que as ideias matemáticas e filosóficas de Turing ainda hoje são o ponto de partida para pesquisas que se desenvolvem na Ciência da Computação e na Inteligência Artificial. Turing acreditava que não havia razão para colocar limites aos progressos que possam ser alcançados nessas áreas. Ele afirmava, contra Descartes, que se para pensar é preciso ter um espírito ou uma alma, e se só Deus pode dar essa alma, não há nenhuma razão que o impeça de fornecer uma alma para um computador. Afinal de contas, Deus é onipotente e dá almas para quem Ele desejar.

Turing dizia que a questão "Pode uma máquina pensar?" só poderia ser respondida no dia em que chegássemos a uma definição precisa dos conceitos de "pensamento" e de "máquina". Esse dia ainda está longe de chegar e, nesse caso, o que faremos até lá?

Se para estabelecermos "o que pensa" e "o que não pensa" devemos nos basear em juízos de senso comum, poderíamos encomendar uma pesquisa de opinião para o Instituto Datafolha. A atribuição de pensamento a organismos ou máquinas seria apenas uma questão social e estatística. Se a maioria das pessoas achar que máquinas pensam, só nos resta aceitar essa opinião, pelo menos provisoriamente. Outra alternativa, proposta por Turing, é substituir essa questão por aquilo que ele chamou de "Jogo da Imitação" e que mais tarde passou a ser conhecido como "Teste de Turing".

O Teste de Turing baseia-se na seguinte ideia: se uma máquina apresentar um comportamento exatamente igual ao de um ser humano, não há por que não atribuir a ela pensamentos e estados mentais. O Jogo da Imitação, que foi a primeira versão do Teste de Turing, requer a participação de três pessoas: um homem (A), uma mulher (B) e um interrogador (C), o qual poderá ser de qualquer um dos sexos. O interrogador permanece num quarto fechado, separado dos outros dois. O objetivo do jogo, para o interrogador, é determinar qual é o homem e qual é a mulher. O interrogador não pode ver "A" ou "B" nem tampouco ouvir suas vozes, caso contrário o jogo terminaria imediatamente. Ele pode se comunicar com "A" e com "B" por meio de uma tela e um teclado de microcomputador. "C", o interrogador, deve fazer uma série de perguntas para "A" e para "B" com o objetivo de determinar seu sexo. "A" e "B" tentarão sistematicamente ludibriar o interrogador.

As perguntas que "C" faz para "A" e para "B" podem ser do seguinte tipo:

C-Qual é o comprimento do seu cabelo?

A ou B-Meu cabelo é curto, os fios mais longos não passam de vinte centímetros.

Um homem normalmente tem cabelos curtos, mas uma mulher também pode usá-los. As respostas de "A" e de "B" serão sempre esquivas. Se ocorre uma situação na qual, por exemplo, "A" seja obrigado a mentir para evitar que o jogo acabe, "B" pode interferir para tumultuar a atividade. "B" poderia dizer, por exemplo: "Não dê atenção ao que ele está dizendo, ele quer ludibriá-lo".

Suponhamos agora que em vez de colocarmos um ser humano no lugar de "A" ou de "B" coloquemos uma máquina. Se no final do jogo "C" não descobrir que estava dialogando com uma máquina e não com um ser humano, podemos afirmar que essa máquina passou no Teste de Turing, ou seja, o comportamento da máquina tornou-se

indistinguível do de um ser humano. Seria legítimo atribuir a essa máquina pensamentos e estados mentais, pois, se assumimos que seres humanos pensam, que outro critério haveria para atribuir pensamentos a algo, além da verificação de que imita perfeitamente o comportamento humano?

O Teste de Turing levanta a seguinte questão: por caso as máquinas não poderiam realizar algo que deveria ser descrito como pensamento, mas que é muito diferente do que um homem faz? Isto significaria dizer que não só o pensamento não é exclusividade dos seres humanos como também que um sistema artificial, construído com material muito diferente daquele que compõe os seres vivos, seria capaz de pensar, desde que tal sistema possa executar as mesmas tarefas que um ser humano.

Foi este o grande projeto teórico da Inteligência Artificial: construir máquinas e programas computacionais que pudessem imitar as atividades mentais humanas. Na década de 1960 e no início da década de 1970 espalha-se uma grande euforia pela Europa e pelos Estados Unidos com a exibição de máquinas que podem demonstrar teoremas automaticamente, realizar cálculos altamente complexos ou jogar xadrez muito melhor do que alguns campeões mundiais. Desse grande entusiasmo com as realizações da Inteligência Artificial surge uma nova teoria da mente, batizada com o nome de *funcionalismo*.

O funcionalismo pretendia se uma teoria completa e abrangente, que explicasse não só das atividades mentais humanas como também as desenvolvidas por sistemas artificiais. Os funcionalistas sustentavam que o mental é o resultado da capacidade de um organismo ou sistema realizar certas funções. Nosso cérebro é um sistema que pode realizar algumas atividades que produzem aquilo que chamamos de mente, mas um outro sistema, como por exemplo um computador, construído com materiais inteiramente distintos, como silício e cobre, será capaz de

produzir atividade mental se ele puder desempenhar as mesmas funções realizadas pelo cérebro humano. Pensar é desempenhar um conjunto de funções que frequentemente levam à produção daquilo que chamamos de comportamento inteligente. Vemos por aí o quanto os funcionalistas estavam influenciados pela concepção de pensamento defendida por Turing!

Mas será correta essa maneira de caracterizar a natureza do pensamento? Até que ponto podemos confiar no Teste de Turing como critério para decidir entre o que pensa e o que não pensa?

A aplicação do Teste de Turing pode ser bastante sugestiva. Basta lembrarmos, por exemplo, de uma das cenas principais do filme de Ridley Scott, *Blade Runner: o caçador de androides*. Nessa cena, o funcionário de uma fábrica especializada na construção de robôs aplica uma longa bateria de questões a um indivíduo para descobrir se ele é um androide ou um ser humano. O objetivo do questionário é bastante preciso; trata-se de saber se esse indivíduo será capaz de responder as questões satisfatoriamente, ou seja, respondê-las da mesma maneira que um ser humano o faria. Se isso acontecer, então saberemos que ele é autêntico ser humano. Todos os seres humanos passariam, em princípio, no Teste de Turing. Se esse indivíduo fosse um androide tão perfeito que pudesse responder com sucesso a todas as perguntas, então ele seria indistinguível de um ser humano. Mas haverá um limite para o número de questões que podem ser formuladas no Teste de Turing ou no Jogo da Imitação? Quando devemos parar? E se o número de questões se tornar infinito, não seria o mesmo que dizer que o teste é inconclusivo?

Mas seria essa a única objeção que poderíamos fazer ao teste inventado por Turing e sua concepção da natureza do pensamento? Suponhamos que você tenha um vizinho que toque piano esplendidamente. Um dia você passa pela sua rua e ouve sons de um piano,



magnificamente executados. Você quer entrar para cumprimentá-lo, bate na porta e aí verifica que não havia ninguém tocando piano, era um aparelho de CD ligado. Contudo, a imitação era perfeita, produziu-se um comportamento indistinguível daquele de um ser humano que sabe tocar piano. As principais condições requeridas pelo Jogo da imitação foram respeitadas: por exemplo, eu passava pela rua e, portanto, não podia ver se quem tocava era um ser humano ou uma máquina. Ora, se Turing estiver certo, seria legítimo atribuir pensamento ao aparelho de CD. Mas terá isto sentido?

Esta estranha conclusão deve-se ao fato de que a comparação entre o homem e máquina estabelecida pelo Teste de Turing recai sobre atividade específicas. Comparam-se apenas alguns segmentos isolados de comportamento humano com aquilo que algumas máquinas ou programas computacionais podem fazer. Se admitirmos esse pressuposto sem crítica, teremos de atribuir pensamento a calculadoras de bolso ou a outros computadores que executam tarefas específicas, como, por exemplo, processar a folha de pagamento do INSS do Brasil inteiro. Certamente a máquina pode executar tal processo com muito mais perfeição e rapidez do que nós, e, nesse caso, teríamos de atribuir a ela uma capacidade de pensamento e raciocínio muito maior do que aquela de qualquer ser humano.

Ora, esse paradoxo surge da identificação de pensamento com a realização de tarefas consideradas de modo isolado. Mesmo quando se aplica o Teste de Turing ao androide de **Blade Runner** ocorre a escolha de capacidades específicas como a memória e a expressão lingüística. Seria preciso comparar as atividades do ser humano como um todo com aquilo que uma máquina pode fazer. Se seguirmos essa linha de raciocínio, veremos que o que caracteriza um ser pensante não é a capacidade de realizar esta ou aquela tarefa, mas precisamente a

possibilidade de variar seus comportamentos. Mas aqui também podemos nos deparar com conclusões tão singulares como: se uma máquina puder variar seu comportamento e fazer tudo aquilo que um ser humano faz, por que continuar a considerá-la uma máquina? E, nesse caso, que sentido teria continuar a indagar se uma máquina pode pensar? Por outro lado, se admitirmos que o que caracteriza o pensamento não é a realização de tarefas específicas, então deveremos concluir, com base no exemplo anterior, que para calcular toda a folha de pagamento do INSS *não é preciso pensar*.

Esses paradoxos mostram como é difícil traçar uma linha nítida que separe aquilo que se deve ser considerado pensante e aqueles artefatos e criaturas aos quais não seria legítimo atribuir pensamentos. A natureza daquilo que chamamos de pensamento ainda permanece misteriosa, e a invenção do Teste de Turing não parece ter contribuído muito para desvendar esse enigma. Suponha que um dia você abra o jornal e leia a seguinte manchete: "ROBÔ JAPONÊS MATA OPERÁRIO NUMA FÁBRICA". Terá esta notícia algumnexo? Certamente não.

Uma parte do robô pode ter escapado ao controle e causado a morte do operário. Mas isso é muito diferente da informação de que o robô matou o operário. O robô pode mover os seus braços mecânicos e executar movimentos que sejam fatais para alguém que esteja por perto, mas isto não significa matar alguém. Somos nós que interpretamos os movimentos do robô como se eles fossem comportamento humano e, ao fazê-lo, inconscientemente atribuímos pensamentos e intenções à máquina. Uma ação envolve algo mais do que a simples observação de suas características externas e, quando as descrevemos, implicitamente atribuímos diferentes tipos de estados mentais que a produziram. É isto que distingue um crime de um acidente; podemos especificar esses tipos de comportamento pela própria descrição que deles fazemos,

embora do ponto de vista de quem aplica o Teste de Turing não haja nenhuma distinção. Turing possivelmente diria que isto pouco importa pois, tanto criminoso como aquele que se envolve em um acidente, são criaturas que pensam. Mas se tudo depende de descrições e se o ponto de partida do teste é a atribuição de pensamento a sistemas ou criaturas - uma atribuição que é feita por aquele que aplica o teste - não estaremos diante de um imenso círculo vicioso onde se pressupõe o pensamento para depois tentar identificar o que pensa e o que não pensa? Nesse caso não teremos feito nenhum progresso na tentativa de explicar a natureza do pensamento.

Esse enigma torna-se ainda mais complicado se associamos ao pensamento a noção de *consciência*. Descrevemos um comportamento consciente como aquele no qual sabemos o que estamos fazendo. Existirá pensamento sem consciência? Os psicanalistas sempre se esforçaram em nos convencer de que é possível dissociar essas duas noções, mas mesmo eles reconhecem que boa parte de nossos pensamentos é consciente. Se essa última afirmação é verdadeira, então fica ainda mais difícil imaginar como uma máquina que pensa poderia adquirir consciência de seus pensamentos.

Alguns filósofos da mente diriam que sem consciência não há sequer ação, quanto mais pensamento. Uma máquina de calcular não faz somas, ela apenas imprime sinais no seu visor de cristal ou sua fita de papel. Para que ela fizesse uma soma, ou seja, estivesse de fato executando uma ação, a máquina precisaria saber, passo a passo, aquilo que está executando. Isto não seria possível se ela não fosse consciente.

Mas será que aqui também podemos traçar uma linha divisória nítida que separe criaturas conscientes e inconscientes? Há muitas ações que executamos de maneira inconsciente e nem por isso deixamos de nos considerar criaturas conscientes. O que dizer de ações que são

executadas durante ataques de sonambulismo ou durante sessões de hipnotismo? Não serão esses estados semelhantes àqueles de um ator que, quando recita seu *script*, deixou de ser uma criatura consciente enquanto vivia exclusivamente uma personagem? E, nesse caso, será que podemos dizer que o sonâmbulo, o hipnotizado e o ator deixaram de ser criaturas conscientes? Da mesma maneira, como poderíamos saber se uma máquina subitamente adquiriu consciência de seus pensamentos e tornou-se uma criatura consciente, como ocorre com o computador HAL do filme *2001- Uma odisseia no espaço*?

Essas dificuldades parecem apontar para a dimensão dos problemas com os quais estamos lidando aqui: pouco sabemos sobre a natureza do pensamento e muito menos acerca do que é consciência. Estes são dois imensos problemas com os quais a Filosofia da Mente tem de se confrontar e sobre os quais ainda não foi capaz de produzir teorias conclusivas. Mas se não sabemos o que é pensamento nem tampouco o que é consciência isto não significa que tentar abordar tais problemas mediante a investigação em Inteligência Artificial seja uma tarefa fadada ao fracasso. É muito difícil profetizar negativamente sobre aquilo que a ciência e a tecnologia poderão vir a descobrir algum dia. Só para se ter uma ideia, basta lembrar que, no início do século XX, velhas senhoras que brandiam seus guarda-chuvas e até mesmo um respeitável filósofo austríaco – Ludwig Wittgenstein - diziam com todas as letras que o homem jamais poderia chegar à Lua.

# A IDENTIDADE PESSOAL

---

Suponhamos que um dia você acorde e note que seus pés estão muitos distantes de sua cabeça. Para calçar suas meias e seus sapatos você precisa esticar os braços muito mais do que estava acostumado a fazer. Ao caminhar em direção ao banheiro, você esbarra a cabeça no topo da porta. Há uma sensação meio esquisita. Quando se aproxima do espelho para se barbear, você leva um tremendo susto: verifica que aquele não é seu rosto e que seu corpo está inteiramente diferente. Ao lado do espelho, há um bilhete pendurado na parede: "Não se assuste. A noite passada, enquanto você dormia, transplantamos o seu cérebro para o corpo do seu vizinho de cima".

Por que o susto? Esta é com certeza uma situação imaginária, mas não impossível. Christian Barnard, o grande cirurgião sul-africano que realizou o primeiro transplante de coração, declarou para a revista *Newsweek*, em 1968, que transplantes de cérebro poderiam ser realizados num futuro próximo. Se isso acontecer, a situação descrita acima não deveria causar tanto espanto.

Mas por que a estranheza, então? Ela surge porque esse tipo de situação nos obriga a refletir sobre a seguinte questão: O que faz com que eu seja eu? Ou, em outras palavras, o que me confere *identidade pessoal*? Será meu corpo, será meu cérebro (que nada mais é do que uma parte do meu corpo), ou será minha mente? E o que faz com que eu tenha certeza de que a criatura que hoje almoça e anda na rua é a mesma criatura que ontem executou essas mesmas ações, ou seja, eu mesmo?

A questão da natureza da identidade pessoal é crucial para a Filosofia da Mente. Sem ela não haveria subjetividade nem tampouco esperança de vida após a morte. Se não existe um princípio que unifica os pensamentos e sensações como sendo *meus* pensamentos e

*minhas* sensações, não há como falar em subjetividade. E se sabemos que o corpo se desintegra após a morte, como poderíamos supor, sem esse princípio unificador, que *pessoas* sobrevivem após a morte?

Essas indagações sempre preocuparam os filósofos e os levaram a esboçar diversas teorias acerca da natureza da identidade pessoal. Teorias clássicas, como a do filósofo inglês John Locke (1632-1704), sustentavam que a identidade pessoal surge da existência de uma consciência que acompanha e unifica as diversas experiências desde o passado até o presente. Minha identidade pessoal dependeria de minha memória e da capacidade que ela tem de organizar uma história de vida que deveria ser sempre única, peculiar. Minha memória possibilita tal consciência unificadora e esta, por sua vez, seria garantia de que a pessoa que levantou ontem e caminhou pelo parque é a mesma pessoa que se levanta hoje e sai para trabalhar. Com isso estaríamos pelo menos garantindo um primeiro tipo de identidade pessoal, ou seja, a identidade ao longo do tempo. Memória e continuidade no tempo: estes seriam os pilares de qualquer identidade pessoal possível.

Mas serão a continuidade de memória e a organização de minhas lembranças suficientes para garantir minha identidade pessoal? Uma primeira objeção a este ponto de vista seriam os casos de amnésia total. Se devido a algum acidente eu perder todas as minhas memórias sobre meu passado, isto significaria que daquela data em diante eu passaria a ser *outra* pessoa? Os defensores da teoria da memória diriam: certamente você passará a ser *outra pessoa*. Mas, e meu eu anterior, terá ele sobrevivido? E quando meu corpo se desintegrar, haverá dois eus possíveis que correspondiam a ele? Esta seria sem dúvida uma dificuldade considerável a ser enfrentada por uma teoria como a de Locke.

Mas não é preciso ir tão longe como nos casos de amnésia total. Basta supor que minha memória pode falhar

e que é possível que eu não me lembre de algumas ações que executei no passado, da mesma maneira que não me lembro da cor da camisa que vesti ontem na hora do almoço. Será então que essas ações foram executadas por mim ou por *outra* pessoa? E se eu tomar um porre fenomenal e sair pelas ruas fazendo coisas terríveis, e não me lembre de ter feito nenhuma delas no dia seguinte? Estas eram objeções que se faziam à teoria de Locke, já na sua época, principalmente por filósofos como Joseph Butler (1692-1752) e Thomas Reid (1710-1796). Não basta ter lembranças e organizá-las numa seqüência temporal, é preciso identificar o princípio que as torna *minhas* lembranças; só assim poderei incorporar ao meu eu coisas que fiz, mas das quais não me lembro ou não posso me lembrar.

Identificar esse princípio pode ser uma tarefa particularmente árdua. Claro que na época de Locke e de Butler não havia robôs, sistemas artificiais e tampouco se sonhava com a possibilidade de realizar um transplante de cérebro. Mas isso não nos impede de imaginar um experimento, utilizando robôs e andróides, no qual possamos testar a validade da teoria da identidade pessoal baseada na memória ou pelos menos explicitar melhor as suas dificuldades.

Suponhamos que um dia alguém construa um sistema artificial, um andróide ou uma réplica perfeita de um ser humano que receba todas as minhas memórias. Será que poderíamos dizer que esse andróide é exatamente como eu ou que tem a minha identidade? Receber um conjunto de memórias, implantadas artificialmente, não é suficiente para que esse andróide as reconheça como sendo *suas* memórias, e é isto que não permite que ele adquira minha identidade pessoal ou outra qualquer. Ele receberá essas memórias de modo impessoal, pois *serem minhas* ou *serem dele* não parece ser uma característica intrínseca de nenhum tipo de estado mental. O andróide não será como

eu; com isto garanto minha identidade pessoal, mas também verifico que o critério de continuidade de memória teria de ser abandonado. Dizer que esse critério precisaria ser complementado por algum princípio que torna um conjunto de memórias *minhas* memórias não explica a natureza da identidade pessoal. Ao contrário, ele simplesmente a pressupõe como algo ainda mais básico e anterior à ideia de que ela deve se basear na consciência de uma continuidade dessas lembranças ao longo do tempo. Talvez as teorias clássicas da identidade pessoal devam simplesmente ser abandonadas.

O que fazer então? Como poderíamos fundamentar a noção de identidade pessoal? Voltemos à situação inicial do transplante de cérebro que descrevemos no início deste capítulo. Ela sugeria que só existem três alternativas nas quais podemos basear o princípio de identidade pessoal: (1) ela teria de ser conferida pelo meu corpo (2) pela minha mente ou (3) pelo meu cérebro. Mas essas três alternativas levam a grandes dificuldades na medida em que não conseguimos até agora resolver uma das questões cruciais da Filosofia da Mente, qual seja, o problema mente-cérebro.

Se abandonarmos as teorias clássicas como aquela proposta por Locke e ainda assim supusermos que a mente é responsável pela identidade pessoal, não há como explicar, na situação do transplante, por que eu sentiria tanta estranheza ao olhar no espelho e ver um outro corpo lá refletido. Essa estranheza não deveria ocorrer, pois, afinal, o corpo seria apenas o invólucro e o suporte biológico da minha mente. E de que adiantaria eu saber que continuei a ser eu mesmo, apesar de ter o corpo de meu vizinho, se todo mundo acreditasse e me tratasse como sendo meu vizinho? Eu teria de provar para essas pessoas que eu sou eu e não meu vizinho, mas como poderia encontrar um meio de fazer tal demonstração se não tenho conhecimento preciso daquilo que confere minha própria



identidade pessoal? Qualquer habilidade que fosse exclusivamente minha ou característica distintiva de personalidade poderia ser aprendida ou desenvolvida por outra pessoa.

Por outro lado, ao afirmar que meu corpo é responsável pela minha identidade pessoal, isto significaria que, se perdesse alguma parte dele, eu deveria, ao mesmo tempo, perder uma parte da minha identidade pessoal. Se perco uma perna ou um braço num acidente, uma parte de meu "eu" também iria junto. Mas é um absurdo pensar que deixo de ser eu por não ter uma perna ou um braço, ou que continuo a ser eu apenas parcialmente. Mas o que aconteceria se todas as partes de meu corpo fossem substituídas? Estaríamos diante de um paradoxo igual àquele do navio de Teseu: quantas tábuas precisam ser substituídas nesse navio para que ele deixe de ser o navio de Teseu como era originalmente? Um navio com algumas tábuas substituídas continua sendo o mesmo navio, mas o que dizer se todas elas fossem substituídas?

A essa altura alguém poderia dizer: você pode perder todas as partes de seu corpo, elas poderiam ser todas substituídas por equivalentes mecânicos, mas você não deixaria de ser você se o seu cérebro não fosse substituído. Ele seria uma parte especial, privilegiada, de meu corpo, o "lugar" da minha mente e a sede de minha própria identidade pessoal. Contudo, uma nova série de paradoxos pode surgir. Se suponho que minha mente e meu cérebro são a mesma coisa e que este último é responsável pela minha identidade pessoal, quando digo "eu sou o JFT" isso corresponde a um estado cerebral. Ora, quem me garante que tal estado não possa ocorrer *em alguém mais* que tenha um estado cerebral idêntico ao meu (talvez como resultado das mesmas reações químicas) e que esse alguém não saia por aí dizendo "Eu sou o JFT"? E quem me garante que esse alguém, por ter um cérebro idêntico ao meu, não pode ter todas as vivências e lembranças que eu tenho, como

resultado de um conjunto de reações químicas idênticas àquelas que ocorrem no meu cérebro? E, nesse caso, quem será o JFT, ele ou eu? Encontramos aqui mais um paradoxo. Se defendemos o materialismo, torna-se impossível definir a identidade pessoal. A única alternativa que resta é supor então que minha mente é responsável pela minha identidade pessoal, pois só assim a sentença "Eu sou o JFT" poderia expressar uma realidade singular, única no mundo. Mas se tenho de supor uma distinção entre mente e cérebro e que esta última é a base da identidade pessoal, isto significa que até agora estivemos girando em círculos...

Essas dificuldades levaram a Filosofia da Mente contemporânea a tentar se apoiar nos resultados da neurociência para superar os paradoxos que envolvem o problema da identidade pessoal. Algumas pesquisas foram aquelas feitas a partir das chamadas *comissurotomias* do cérebro. A comissurotomia ou bissecção cerebral consistia em separar os dois hemisférios de cérebro cortando cuidadosamente o conjunto de ligações nervosas que os une, o chamado *corpus callosum*. Essas operações foram realizadas principalmente no fim dos anos 1940 e tinham por objetivo proporcionar algum alívio a pacientes acometidos de fortes ataques de epilepsia.

Ocorre que as comissurotomias levaram a resultados particularmente interessantes. Foi constatado que os dois hemisférios podem funcionar de maneira quase independente e que o hemisfério esquerdo é o responsável pela produção da linguagem. Foi possível isolar, em situações especiais, os dois hemisférios e submetê-los, separadamente, a um mesmo tipo de experiência. Verificou-se então que eles poderiam reagir de maneira diferente sem, entretanto, se comunicar entre si.

Os neurocientistas viveram em período de grande entusiasmo nessa década. Eles achavam que com seus experimentos estavam a caminho de encontrar uma

resposta definitiva para as principais questões acerca da natureza da identidade pessoal que normalmente preocupavam os filósofos. Afinal, não estariam eles próximos de localizar o lugar do "eu" no mapa neurológico, a verdadeira sede da consciência e da identidade pessoal? Não seria razoável supor que este "eu" estaria localizado do lado esquerdo do cérebro, o qual, uma vez sendo responsável pela produção da linguagem, seria também o produtor da consciência? Afinal, delineava-se a possibilidade de formular, coerentemente, uma teoria científica e materialista da natureza da identidade pessoal.

Entretanto, esse entusiasmo inicial foi logo corroído pelo aparecimento de mais dúvidas e indagações que surgiram na medida em que se discutia a interpretação dos resultados desses experimentos. Se os dois hemisférios poderiam reagir de maneira independente e sem se comunicar, quantas mentes ou quantos "eus" teria esse paciente submetido a comissurotomia? Haveria uma pessoa ou duas pessoas? Será que esses pacientes não teriam duas mentes, uma que pode falar e outra que não pode? E se a consciência e o "eu" nada mais são do que um conjunto de neurônios, não seria razoável supor que, quando eu falo, não sou eu que estou falando, mas apenas um conjunto de neurônios no meu cérebro, responsável pela produção da linguagem? E como atribuir a eles subjetividade e identidade pessoal? "Não fui eu quem fez isso, foi minha mão", é assim que as crianças falam nas suas brincadeiras; o mesmo teríamos que aplicar a nós mesmos e nos conceber apenas como um feixe de neurônios, cada um executando, independentemente, uma tarefa. Teríamos muitos eus, muitas vozes falando na nossa cabeça, mas a nenhum desses feixes de neurônios poderíamos nos referir como constituindo uma *pessoa*.

A comissurotomia não resolveu o mistério da identidade pessoal, apenas multiplicou os "eus", tornando essa questão ainda mais complexa. Ora, não seria o caso de

dar razão ao poeta português Fernando Pessoa - que era também Alberto Caeiro, Ricardo Reis e Álvaro de Campos - e que no fim de sua vida escreveu em uma de suas odes: "se Deus é um só, por que haveria eu de ser tantos? "

# CONCLUSÃO

---

Teremos percorrido todos os problemas da Filosofia da Mente? Certamente não. Os três problemas de que tratamos aqui - as relações mente-cérebro, a questão da natureza do pensamento e da consciência e o problema da identidade pessoal- são fundamentais, mas não recobrem todas as preocupações dos filósofos da mente. Outros enigmas ainda estão à espera de uma possível solução.

Seria preciso discutir, por exemplo, as relações entre pensamento e linguagem que levaram alguns filósofos, na década passada, a supor que a Filosofia da Mente poderia ser totalmente absorvida pela Filosofia da Linguagem. Mas esta foi uma perspectiva que não resistiu muito tempo às críticas. Sabemos que normalmente pensamos com palavras e que a linguagem é constitutiva do pensamento, mas seria exagero afirmar que a linguagem e pensamento são a mesma coisa. Sustentar esse ponto de vista seria o mesmo que negar a muitos animais e capacidade de pensar, o que certamente seria contra intuitivo. Apesar de haver evidências de que alguns animais têm algo parecido com uma linguagem, esse tipo de código ainda está longe de ser tão bem estruturada quanto o nosso.

Outra questão que mereceria ser discutida e quem tem ocupado os filósofos da mente é o chamado problema da *representação mental*. Como é possível que nossas ideias espelhem o mundo que está fora de nós? Se enxergamos uma vaca amarela no pasto e, se alguém abre nosso cérebro exatamente neste momento, certamente não verá nada amarelo no seu interior. Como pôde então se formar uma imagem de algo amarelo em nossa mente? Ou melhor: como podem se transformar os impulsos elétricos no nosso cérebro em representações de coisas diversas que vemos no mundo? Nosso cérebro tem bilhões de células nervosas e bilhões de ligações entre elas, mas, embora esse número seja colossal, ele certamente é finito. Como podemos

formar na nossa cabeça uma ideia de *infinito*, se nem o nosso cérebro nem as coisas que vemos a nossa volta são infinitas?

Estes são problemas de difícil solução. Da mesma maneira, um desfecho satisfatório para as questões que envolvem as relações mente-cérebro e a natureza da identidade pessoal parece ainda muito distante. Isto levou os filósofos a acharem, em um certo momento, que tais problemas deveriam ser abandonados, pois seriam insolúveis, para além das capacidades da razão humana. Alguns chegaram a afirmar que para resolver o problema das relações mente-cérebro precisaríamos deixar de ser humanos, pelo menos por alguns instantes, e poder contemplar nossa espécie de uma perspectiva divina, supra-humana, uma espécie de cadeira cósmica. Só assim teríamos (pelo menos por alguns momentos) os olhos de Deus e só com eles poderíamos verificar se mente e cérebro seriam ou não a mesma coisa.

Por outro lado, seria possível pensar que a Física pudesse um dia nos ajudar a encontrar um caminho para resolver esse problema. Após as revoluções teóricas realizadas por James Clerk Maxwell (1831-1879) e Albert Einstein (1879-1955), chegou-se à conclusão de que matéria e energia são manifestações de um único tipo de princípio que rege o universo. Não existiria mais nenhuma razão para definir a matéria como algo sólido e visível; ela seria apenas um estado particular da energia. Se matéria e energia são uma única e mesma coisa, só aparentemente distintas, não poderíamos dizer o mesmo da oposição mente/corpo, identificando o mental com um tipo qualquer de energia? Não estaria assim resolvido um problema secular? Esta é sem dúvida uma tentação muito grande, que abriria a perspectiva de resolver, por meio da ciência, um problema filosófico particularmente árduo. Mas não estaríamos aqui correndo o risco de incorrer num círculo vicioso ao tentar

usar produtos da nossa mente - como é a ciência - para explicar a própria natureza do mental?

Essas dificuldades parecem causar em nós uma grande sensação de desânimo. Mas por que a humanidade continua a se colocar essas questões e por que elas voltam a aparecer na filosofia, mesmo que muitos já tenham demonstrado sua insolubilidade? Parece que este é o verdadeiro mistério a ser desvendado. A história da filosofia oscila como um pêndulo, que se alterna entre otimismo e pessimismo, entre metafísica e história natural. É preciso que entre essas oscilações de pêndulo a besta humana não sucumba a um fardo maior que aquele que ela pode carregar. Ou, talvez, que deixemos de ser bestas, esqueçamos tudo isto e simplesmente levemos nosso cérebro para dar uma volta no centro da cidade para tomar um pouco de ar fresco.