



SÉRIE  
SÉRIES  
CONTEMPORÂNEAS

FILO  
CONTE



# INTRODUÇÃO À FILOSOFIA MENTE

2<sup>a</sup> edição

K. T. MASUN

Editora da UNESP  
São Paulo, São Paulo  
www.unesp.br/editora



ISBN  
978-85-352-0000-0

1.000 páginas



**Introdução à Filosofia da Mente** apresenta os principais assuntos contemporâneos, temas e argumentos que atualmente estão sendo debatidos nessa área. O livro examina criticamente quais implicações teóricas do raciocínio dualista, identidade materialista, behaviorismo e functionalismo. Argumenta que enunciados de consciência e memória só fazem sentido dependendo de processos. Tais noções não são resultados e sim processos. As diferenças entre enunciados falsos e mentiras visuais/auditivas mantêm-se, o problema de crenças falsas e a identidade de possibilidades são extensivamente explorados. Esta nova edição inclui um capítulo que explora a Filosofia da psicologia e da mente de Aristóteles, e também um novo material sobre o Teorema de Turing.

Este livro foi concebido para ajudar os estudantes a pensar sobre os temas apresentados e, assim, ao longo do texto, questões e questões para reflexão que estimulem e desafiem o leitor.

## INTRODUÇÃO À FILOSOFIA DA MENTE

Prólogo

S. L. Hirsch



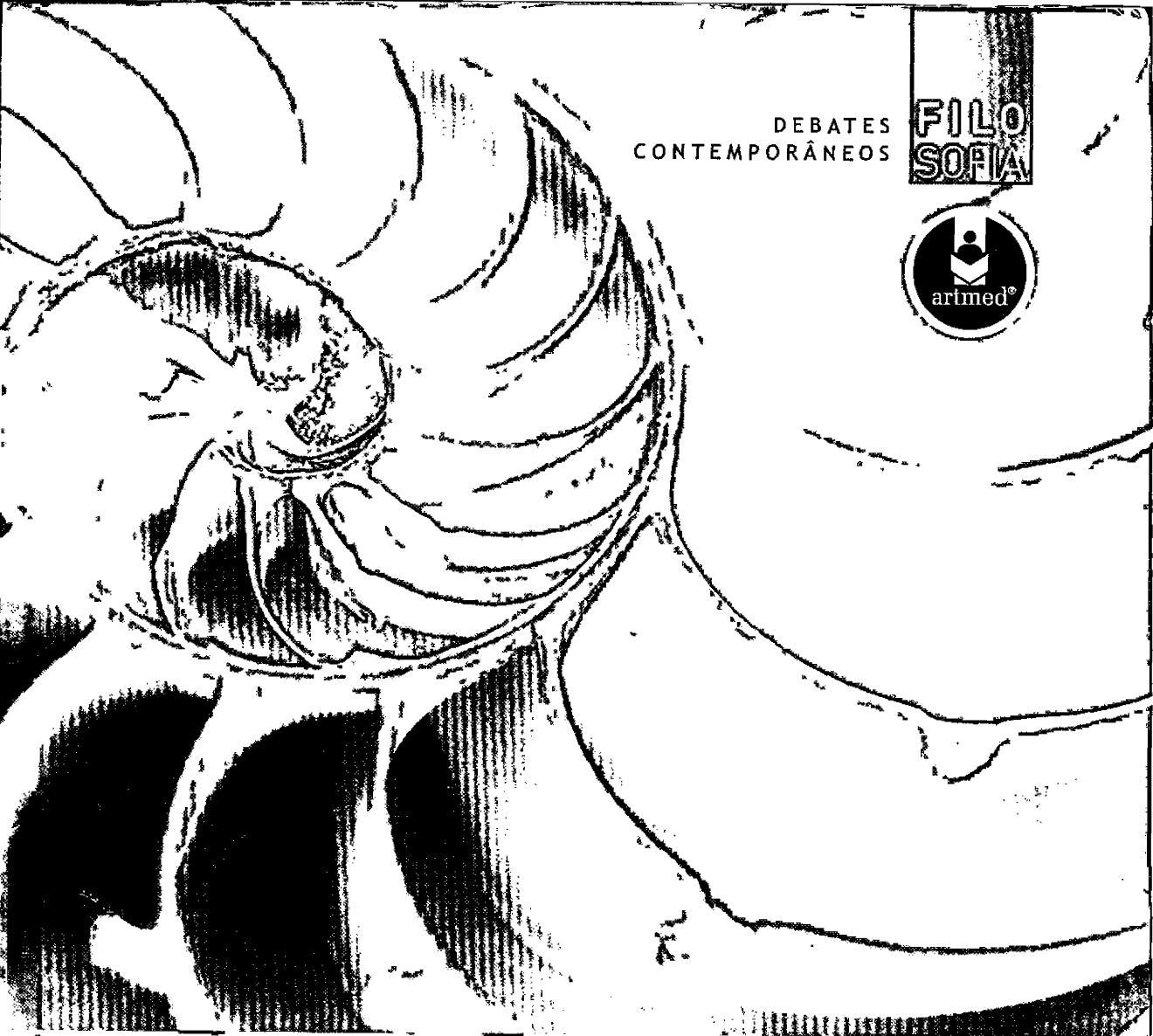
**artmed®**  
EDITORIAL  
MÍDIA E CULTURA



FILO  
SOFIA

INTRODUÇÃO  
À FILOSOFIA  
DA MENTE





DEBATES  
CONTEMPORÂNEOS

FILO  
SOFIA



# INTRODUÇÃO À FILOSOFIA MENTE

2<sup>a</sup> edição

K. T. MASLIN

N.Cham.: 128.2 M397i.2  
Autor: Maslin, K. T.  
Título: Introdução à filosofia da



UFMS CBC

325088  
Ac.118585

## BIBLIOTECA ARTMED FILOSOFIA

- \* BONJOUR, L. – Compêndio de filosofia
- \* BROUGHTON, J.  
Um guia para o estudo de Descartes
- CAREL & GAMEZ  
Filosofia contemporânea em ação
- CUNNINGHAM, F.  
Teoria da democracia: uma introdução crítica
- CUSSET, F. – Filosofia francesa: influência de Foucault, Derrida, Deleuze & Cia
- \* DOSSE, F. – Giles Deleuze/Félix Guattari
- \* FERRY, L. – Filosofia política
- FRENCH, S.  
Ciência: conceitos-chave em filosofia
- FURROW, D.  
Ética: conceitos-chave em filosofia
- GOLDSTEIN, BRENNAN, DEUTSCH & LAU  
Lógica: conceitos-chave em filosofia
- GARRETT, B.  
Metafísica: conceitos-chave em filosofia
- GOMBAY, A. – Descartes
- \* HACKER, P.M.S. – Natureza humana
- HÖFFE, O. – Aristóteles
- \* KRAUT, R.  
Um guia para o estudo da ética em Aristóteles
- MASLIN, K.T. – Introdução à filosofia da mente
- MATTHEWS, E.  
Mente: conceitos-chave em filosofia
- \* MCMAHAN, J. – A ética do ato de matar
- MEDINA, J.  
Linguagem: conceitos-chave em filosofia
- MORRIS, K. – Sartre
- NORRIS, C.  
Epistemologia: conceitos-chave em filosofia
- \* WHIMSTER, S. -- Weber
- WOOD, A.W. – Kant

\* Livros em produção no momento da impressão desta obra, mas que muito em breve estarão à disposição dos leitores de língua portuguesa.

INTRODUÇÃO  
À FILOSOFIA  
DA MENTE





M397i Maslin, K. T.

Introdução à filosofia da mente / K. T. Maslin ; tradução Fernando José R. da Rocha. – 2. ed. – Porto Alegre : Artmed, 2009.

304 p. ; 25 cm.

ISBN 978-85-363-1926-1

1. Filosofia. 2. Filosofia da mente. I. Título.

CDU 13

Catalogação na publicação: Renata de Souza Borges CRB-10/1922

# INTRODUÇÃO À FILOSOFIA DA MENTE

2<sup>a</sup> edição

K. T. MASLIN

Head of Philosophy and Critical  
Thinking at Esher College, Surrey, UK

Consultoria, tradução e supervisão desta edição

Fernando José R. da Rocha  
Professor Associado do Departamento de Filosofia  
da Universidade Federal do Rio Grande do Sul



2009

Obra originalmente publicada sob o título:  
*An Introduction to the Philosophy of Mind*, 2.ed.

ISBN 978-0-7456-4074-7

© 2007, Keith Maslin

This edition is published by arrangement with Polity Press Ltd., Cambridge.

Capa: Tatiana Sperhache - TAT Studio

Foto da capa: © iStockphoto.com/Skodonnell

Preparação do original: Maria Edith Amorim Pacheco

Leitura final: Carlos Henrique Lucas Lima

Supervisão editorial: Mônica Ballejo Canto e Carla Rosa Araujo

Editoração eletrônica: VS Digital

|         |            |
|---------|------------|
| Class.: | 128,2      |
|         | M 3971.i.2 |
| Tombo:  | 325088     |

Reservados todos os direitos de publicação, em língua portuguesa, à  
ARTMED® EDITORA S.A.  
Av. Jerônimo de Ornelas, 670 - Santana  
90040-340 Porto Alegre RS  
Fone (51) 3027-7000 Fax (51) 3027-7070

É proibida a duplicação ou reprodução deste volume, no todo ou em parte,  
sob quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação,  
fotocópia, distribuição na Web e outros), sem permissão expressa da Editora.

SÃO PAULO  
Av. Angélica, 1091 - Higienópolis  
01227-100 São Paulo SP  
Fone (11) 3665-1100 Fax (11) 3667-1333

SAC 0800 703-3444

IMPRESSO NO BRASIL  
PRINTED IN BRAZIL

## AGRADECIMENTOS

---

Gostaria de expressar minha gratidão a um grande número de pessoas, sem as quais este livro jamais poderia ter sido escrito.

Sou grato a Rebecca Harkin, que antes estava na Polity, e agora está na Blackwell, pela encomenda do livro. Foram sua paciência, seus aconselhamentos e seu constante apoio que me fizeram prosseguir, quando de outra forma poderia ter desistido. Também gostaria de agradecer a John Thompson, Louise Knight e Gill Motley, da Polity, Cambridge, por suas valiosas sugestões, e, particularmente, a três referees anônimos, cujos comentários positivos e críticas úteis permitiram-me fazer melhorias substanciais ao texto. Pam Thomas, a gerente editorial da Polity, também me deu muitos conselhos úteis e estímulo; e eu devo agradecer a Kathryn Murphy, a controladora de produção, por todo o duro trabalho que dispensou à produção do livro. Um agradecimento muito especial deve ir à editora para publicação, Sara Dancy, que leu todo o texto com imenso cuidado, e quem não só me possibilitou evitar vários erros, mas também fez excelentes sugestões para tornar o livro mais acessível e de uso mais amistoso.

Em 1996, não muito depois de eu haver começado a escrever o primeiro rascunho, tive a felicidade de receber uma bolsa de estudos, outorgada pelo Farminghouse

Institute of Christian Studies, para um quadrimestre no Harris Manchester College, University of Oxford. Esse período, em que estive livre de obrigações docentes e administrativas no Esher Sixth Form College, foi de grande valor. Devo agradecer primeiramente ao Rt. Hon. Bobby Wills, que criou o Instituto, e, em segundo lugar, a Martin Rogers, OBE, o diretor, por seu continuado estímulo e aconselhamento. Agradeço também a Ron Glithero, que leu e discutiu grande parte da primeira versão, e a Suzanne Tetsell e Elizabeth Machin, que proporcionaram apoio em nível de secretaria e administração de que muito necessitei durante minha estada no Harris Manchester. O Reverendo Dr. Ralph Waller, diretor da faculdade, nos fez sentir muito bem-vindos, mas agradecimentos especiais devem ir a Roger Trigg, professor de filosofia da Warwick University que era um professor visitante durante o período em que estive em Oxford. Roger, muito generosamente, ofereceu-se para ler tudo aquilo que escrevi e posteriormente discutir na santidade das salas de chá de Wykeham. Eu rememoro tais diálogos com enorme satisfação.

No entanto, eu não poderia ter assumido a bolsa de estudos sem o apoio do Dr. Patrick Miller, então diretor do Esher College, e sou igualmente grato ao então vice-diretor, Geoff Mason, por sua disposição em fazer as tratativas necessárias ao

meu afastamento. Também fiquei muito satisfeito que o Dr. Miller pudesse ter assistido à apresentação do trabalho que eu havia completado ao final de minha bolsa de estudos em Oxford. Dr. Nick McAdoo muito generosamente me substituiu durante um quadrimestre enquanto estive afastado de Esher e ajudou a manter o Departamento de Filosofia em ordem.

Não devo esquecer de mencionar os membros do South-West London Gastro-nomico-Philosophical Society, que ouviram e fizeram muitos comentários de utilidade sobre muitos capítulos, assim como um grande número de alunos meus em Esher, com os quais eu testeи grande parte do material que se encontra no livro, e que me trouxeram importante retorno, além de forte estímulo para a conclusão do projeto. Basil Hunt, anteriormente chefe do Departamento de Civilização Clássica no Esher College, também fez sugestões que foram úteis ao aprimoramento do estilo da escrita, bem como estimulantes discussões sobre algumas das principais ideias contidas no livro.

Devo também agradecer a minha esposa, Zielfa, a nossa filha Philippa, e a meu pai, pela tolerância e pelo apoio enquanto o projeto estava em andamento.

Minha dívida para com o trabalho de filósofos desde Descartes até Wittgenstein, bem como para com muitos escritores contemporâneos sobre a filosofia da mente, por demais numerosos para serem mencionados, se tornará óbvia no próprio texto. Não tenho a pretensão de haver feito uma contribuição original à filosofia da mente, mas o livro não é simplesmente um levantamento neutro do assunto, e, em geral, o leitor não deve ficar em

dúvida quanto aos meus próprios pontos de vista sobre vários tópicos ao longo do texto. Não quero criar a impressão de que não restam perguntas e de que todas as dificuldades foram resolvidas, quando claramente não é o caso. Tenho a esperança de que os leitores darão por conta própria continuidade às questões que permanecem em discussão. Acima de tudo, como este é um texto introdutório, tentei escrever de modo claro e acessível, pressupondo o mínimo conhecimento prévio da área, que seja razoável.

Finalmente, tenho um débito especial de gratidão à minha mãe, Dorothy Vera Maslin, que, desde meus primeiros anos, ensinou-me a amar a leitura e a vida da mente e que foi importante suporte ao longo dos anos, especialmente em tempos de dificuldade e desânimo. Ela sempre me estimulou a escrever e teria adorado ver a publicação do livro, caso estivesse viva. Este livro é dedicado à sua memória.

O autor e o editor gostariam de agradecer aos a seguir mencionados pelas permissões de citar materiais com direitos autorais: Faber & Faber Ltda e Harcourt Inc. por um extrato do poema "Morning at the Window", de T. S. Elliot, *Collected Poems 1909-1962*; Penguin, UK, pela permissão de reproduzir um diagrama e fazer uma citação do *The Mechanical Mind*, de Tim Crane (Penguin Books, 1995).

Todos os esforços foram feitos para localizar os detentores de direitos autorais. No entanto, caso algum deles tenha inadvertidamente passado despercebido, os editores terão prazer em fazer os acertos necessários na primeira oportunidade.

## COMO UTILIZAR ESTE LIVRO

---

Este livro apresenta uma série de elementos que auxiliarão o leitor a se situar rapidamente, e atingir uma compreensão sólida e profunda daquilo que é necessário saber.

Um rápido panorama do que você encontrará é apresentado por meio de uma lista de objetivos no início de cada capítulo. Esses objetivos especificam os principais temas que você deve saber para compreender as questões desenvolvidas no decorrer do livro. Assim, esperamos que você obtenha uma valiosa relação de conteúdos.

Em segundo lugar, exercícios que visam auxiliá-lo a pensar de maneira autônoma sobre vários tópicos são integrados ao texto durante a leitura. Os exercícios poderão ser resolvidos sozinho, mas frequentemente seu benefício será maior se você discutir a questão com alguma outra pessoa, conforme a sugestão do próprio exercício. Em qualquer caso, minha maior recomendação é a de que você resista à leitura das respostas que eu dou às questões formuladas, até que você tenha formulado suas próprias reações. Dessa forma, você estará ajudando a fortalecer seus próprios poderes de análise e crítica, bem como evitando se colocar demasia-damente sob a influência ou dependência do que eu tenho a dizer.

Uma forma de evitar ler inadvertidamente o que escrevi seria utilizar um cartão opaco e gradualmente destapar o

texto. Do mesmo modo, um professor que queira gerar uma discussão em um pequeno grupo poderia usar os exercícios como estímulo, antes de solicitar aos estudantes que leiam a discussão sobre os vários tópicos ou antes de proceder à apresentação formal das questões.

Em terceiro lugar, uma seleção de perguntas típicas de examinadores de primeiro e último semestres podem ser encontradas no final de cada capítulo, juntamente com indicações para leituras mais aprofundadas para aqueles que desejam prosseguir por conta própria. Certamente os professores poderiam estabelecer tais ensaios para seus alunos. De forma alternativa, os alunos poderiam tentar planejar ou escrever esses ensaios sozinhos.

Em quarto lugar, embora eu tenha tentado escrever do modo menos técnico possível e tenha, ao longo do texto, apresentado e explicado termos que possam não ser familiares para você, ainda resta o perigo de eu haver feito demasiadas pressuposições. Em razão disso, um glossário de termos filosóficos, cujos significados podem não ser claros, encontra-se no final do livro.

Finalmente, uma lista de referências seleta encontra-se no final do livro para aqueles que tenham interesse em aprofundar o assunto para além das indicações constantes nas listas de leituras ao final de cada capítulo.

# SUMÁRIO

---

|   |     |
|---|-----|
| Agradecimentos .....  | v   |
| Como utilizar este livro .....                                      | vii |
| <b>1.</b> O problema mente/corpo .....                              | 11  |
| <b>2.</b> Dualismo .....  | 39  |
| <b>3.</b> A teoria da identidade mente/cérebro.....                 | 72  |
| <b>4.</b> Behaviorismo analítico .....                              | 105 |
| <b>5.</b> Funcionalismo .....                                       | 128 |
| <b>6.</b> Levando a consciência a sério: monismo não-redutivo ..... | 159 |
| <b>7.</b> Causação psicofísica .....                                | 181 |
| <b>8.</b> Uma perspectiva aristotélica.....                         | 201 |
| <b>9.</b> O problema de outras mentes.....                          | 220 |
| <b>10.</b> Identidade pessoal como continuidade física.....         | 247 |
| <b>11.</b> Identidade pessoal como continuidade psicológica .....   | 264 |
| Glossário .....   | 285 |
| Referências .....   | 295 |
| Índice.....   | 297 |

## O PROBLEMA MENTE/CORPO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Ter uma compreensão preliminar do problema mente/corpo e de como ele surge.
- Ter uma compreensão sobre algumas das principais teorias da mente e do corpo disponíveis.
- Compreender as diferenças entre estados intencionais e não-intencionais.
- Compreender o que significa o qualia associado com alguns tipos de estados mentais.
- Compreender a estrutura das atitudes proposicionais e por que se tem pensado que elas apresentam problemas para teorias materialistas da mente.
- Entender a distinção entre intencionalidade intrínseca e derivativa.
- Compreender o que se quer dizer com perspectivas de primeira e terceira pessoas e por que se tem pensado que estados mentais sejam radicalmente privados a pessoas.
- Compreender por que se pode colocar dúvida na assertiva de que uma pessoa sempre tem acesso privilegiado a seus estados intencionais, assim como conhecimento infalível deles.
- Compreender por que se tem pensado que a existência de perspectivas de primeira e de terceira pessoas apresenta um problema para a obtenção de uma abordagem unitária para a mente.
- Compreender o que se quer dizer com o caráter holístico de estados intencionais e por que a autoconsciência é necessária para que a racionalidade se torne possível.
- Compreender os possíveis critérios para a demarcação entre o mental e o físico.

## Introdução

Quando comecei a escrever este capítulo podia ouvir o ruído do computador a minha frente. Minha gata, aninhada sobre um grande cobertor de plumas dobrado, ronronando prazerosamente e, quando me virava, podia vê-la se limpando. Sentia-me com um pouco de dor e rigidez no meu ombro direito, e ligeiramente cheio, pois não fazia muito tempo que havia almoçado. Fazia uma pausa momentânea em minha escrita para visualizar minha esposa, que estava visitando sua terra natal, as Filipinas; e me lembrei da época em que vivi lá, e fomos às montanhas durante as férias. Também pensava em ir buscá-la no aeroporto daí a uma semana, e vagamente me perguntava se ela precisaria de uma refeição e o que eu deveria cozinar.

Por que forneci a você esse pequeno fragmento de minha história de vida? A razão é lembrar-lhe do fato muito óbvio de que na qualidade de um ser humano eu sou um sujeito consciente que tem experiências, e sou possuidor de uma variedade de estados mentais – ouvir, ver, sentir uma dor ou desconforto, ver nos olhos da mente, lembrar, pensar e questionar –, para mencionar alguns casos de estados mentais, selecionados a partir de uma longa lista de possibilidades. Eu sou o centro de um universo de experiências – e você também é. Animais, pássaros, peixes, insetos e possivelmente até moluscos têm vida mental, não obstante quão subdesenvolvidas e rudimentares algumas dessas mentes possam ser.

Mas, então, o que é ter uma mente? Por um lado, essa pergunta é fácil de se responder. Como a breve incursão na minha história mental no início deste capítulo ajuda a ilustrar, ao falar sobre a mente – a sua mente ou a de outras pessoas – você está descrevendo os tipos de estado ou de acontecimentos que compõem uma

mente – pensamentos, crenças, emoções, sensações, intenções, desejos, percepções, propósitos que auxiliam abarcá-la. Por outro lado, no seu próprio caso, parece impossível chegar mais perto de seus próprios estados mentais, uma vez que estes abrangem literalmente o material de sua experiência, sua vida mental e que, sem você, não existiriam.

Por enquanto, essa caracterização do que geralmente chamarei de "fenômenos mentais" é satisfatória. Mas, no final, ela deixa de ser assim, porque ela parece nos dar uma ideia muito pequena, se é que dá alguma ideia, da natureza dos estados mentais. Fazendo um contraste, considere os fenômenos físicos. Existem teorias bem desenvolvidas sobre a natureza e o comportamento de corpos físicos, processos e eventos, como as das ciências, que a química, a física e a biologia exemplificam. Obviamente, há muitas coisas que ainda não conhecemos sobre a derradeira natureza da realidade física, mas parece razoável dizer que temos uma melhor apreensão do que está acontecendo quando se trata de uma ocorrência física que de estados mentais. Por exemplo, a investigação científica está, cada vez mais, descobrindo sobre os processos que ocorrem em nosso corpo e cérebro, e embora a nossa compreensão esteja longe de ser completa, temos ao menos uma boa ideia dos tipos de detalhes que ainda temos de descobrir e como planejar tal tarefa.

Fazendo um contraste, alcançar uma compreensão da natureza de estados mentais não parece poder ser comparado com o que implica atingir uma melhor compreensão de estados físicos. Consequentemente, enquanto todos nós sabemos o que seja debruçar-se sobre um problema intelectual, tal como resolver uma equação algébrica, ou, no outro extremo, sentir uma dor, nossa experiência parece não nos fornecer indicação alguma sobre o que

seja a natureza intrínseca do que ocorre dentro de nós quando pensamos, ou quando uma parte de nosso corpo dói. Temos de admitir que um engajamento ativo na busca de uma solução para um problema em matemática, por um lado, e a experiência passiva de uma sensação de dor, por outro, são tipos extremamente diferentes de fenômenos mentais. Tão disparatados, de fato, que até se hesita em mencioná-los de uma só vez. No entanto, embora sejam muito distintos quanto a seu caráter, tanto o pensar quanto o sentir dor pertencem ao lado mental, em oposição ao lado físico da divisória. Dito isto, no entanto, ainda ficamos com a questão de saber precisamente o que está se passando em nós quando qualquer um desses dois tipos de eventos mentais ocorre. Em que tipo de meio, quer-se perguntar, o pensar e a experiência de sentir dor ocorrem?

É este tipo de quebra-cabeças que conduz e ajuda na constituição do problema mente/corpo. Três questões, sugiro, fazem uma insistente pressão sobre nós. Em primeiro lugar, qual é a natureza da mente, dos estados mentais e dos eventos em geral? Qual é, fazendo uma colocação mais imponente, o modo de ser deles? Isso é o que os filósofos chamam de a questão ontológica a respeito do status da mente. A palavra "ontologia" é derivada da palavra grega *ontia*, que quer dizer "coisas que existem". Ao perguntar qual é a ontologia dos estados mentais, os filósofos querem saber do que eles consistem, o que realmente os constitui. São eles em última instância de natureza física, à semelhança dos processos envolvidos, por exemplo, na fotossíntese, ou são eles totalmente de caráter imaterial, nada tendo em comum com o mundo físico, como as almas e os espíritos são tradicionalmente concebidos? De modo alternativo, deveria o dilema, físico ou não-físico, que parece se impor a nós aqui, ser rejeitado em fa-

vor de outras abordagens da natureza da mente?

Em segundo lugar, as experiências devem existir como sendo experiências de alguém ou de algum ser. A ideia de que poderia haver experiências soltas e separadas flutuando por aí, como se fossem sem ligação a algum ou outro sujeito, não parece fazer sentido, embora essa questão seja controvertida.<sup>1</sup> Caso estejamos preparados para conceder, ao menos por enquanto, que as experiências devem ter donos, somos naturalmente levados a nos perguntar qual é a natureza dos possuidores de estados mentais – aquelas pessoas, assim como os animais, cujos conjuntos de experiências englobam vários tipos de fenômenos mentais. Esses sujeitos deveriam ser concebidos como se existissem além e acima das próprias experiências, ou como de alguma forma constituídos fora delas, como um quebra-cabeças de montar não é nada além ou acima das peças que entram na sua composição? Por outro lado, se os sujeitos de experiências diferem quanto ao tipo das próprias experiências, e não podem ser reduzidos a elas, tais sujeitos têm um caráter físico ou não-físico? Várias respostas a essas questões nos são dadas por meio das cinco diferentes teorias da relação mente/corpo, que compõem os Capítulos 2 a 6.

Em terceiro lugar, está implícito, nas questões suscitadas pelas duas últimas perguntas, o quebra-cabeças de como os estados mentais e os sujeitos se relacionam com o mundo físico. Podem os fenômenos mentais existir independentemente dos fenômenos físicos, ou dependem destes para a sua existência? Se estados físicos de fato fazem com que estados mentais surjam, como pode isto ocorrer, uma vez que o mental e o físico parecem ser tão distintos um do outro? Este problema me foi trazido à casa com força especial faz alguns anos. Eu estava dando um curso de

filosofia da mente para alunos adultos na Rewley House, em Oxford, e eu disse que gostaria, se possível, de ser capaz de mostrar um cérebro humano de verdade à turma. Eu tinha, certamente, visto cérebros em fotografias e em museus de ciências, mas eu estava intrigado para saber como seria dispor de um para inspecionar em detalhes. As autoridades da Universidade foram além das minhas expectativas: não só eles entregaram um cérebro, mas ele havia sido conservado de tal forma que você podia realmente manuseá-lo e passá-lo aos demais alunos. Isso certamente deu aos membros da turma e a mim mesmo uma estranha sensação ao pensar que estávamos manuseando de forma tão casual o que uma vez esteve no crânio de uma pessoa viva. Ao mesmo tempo, um sentimento de questionamento intenso foi provocado pela reflexão: como é que pode este pedaço de matéria pálida cinza-esbranquiçada, mesmo se estivesse completamente funcionando e conectada com o resto de um ser humano vivo, ser responsável pelo surpreendente tecnicolor invólucro imaginário da consciência? Como poderia esta coisa ser a corporificação de esperanças e sonhos, medos, crenças, intenções, memórias e pensamentos de uma pessoa? Não somente parecia haver um hiato entre, de um lado, os cérebros e os processos que se dão dentro deles, e, de outro, a consciência, mas um golfo intransponível que, devido à sua própria natureza, não poderia ser transposto ou eliminado. O grande filósofo do século XVII, Leibniz (1646-1716) encapsulou este quebra-cabeças quando nos solicitou que imaginássemos uma máquina, cuja estrutura produzisse pensamento, sentimento e percepção, aumentada ao tamanho de um moinho, a fim de que pudéssemos visitar seu interior. O que encontrariamos lá?, perguntou Leibniz. Resposta: "Nada além de partes que empurram e movem umas às outras,

e nunca algo que pudesse explicar a percepção".<sup>2</sup> A máquina que Leibniz tinha em mente era, certamente, o cérebro. O que os cientistas modernos encontram quando investigam seu funcionamento são eventos físicos eletroquímicos, os equivalentes das "partes que empurram e movem umas às outras" de Leibniz. Mas se a neurologia não pode ser invocada para explicar a percepção e a consciência, o que pode?

A resposta, de acordo com o dualista (cujos argumentos formam a primeira grande teoria da relação mente/corpo, que ampliarei e avaliarei no Capítulo 2), é que a maquinaria do cérebro nunca pode, em princípio, explicar a existência de estados mentais. Por outro lado, cada um de nós deve ser pensado como entidades não-físicas, uma alma, cujos estados materiais abarcam nossas vidas mentais. A alma é capaz de existir por si própria, em completa independência de qualquer outra coisa. Durante a vida, ela é de alguma forma ligada ao corpo, afetando-o e sendo afetada por ele, mas, após a morte, ela se separará dele. Embora o dualismo possa ser encontrado ao menos desde os gregos antigos, René Descartes (1596-1650) deu à doutrina uma nova forma diferenciada, e argumentos originais e poderosos a seu favor. Em consequência disso, a teoria é frequentemente referida e discutida sob o título de "dualismo cartesiano" ("cartesiano" é o adjetivo formado a partir de seu nome próprio: Descartes).

Para o materialista estrito, a noção de alma é sem sentido, e ele a descarta, insistindo fortemente que a mente não é nada mais do que o cérebro em funcionamento. Os processos mentais não existem acima e além dos processos físicos: eles são processos físicos, e nada além de processos físicos. Ao explicarmos como o cérebro funciona, estamos, de fato, explicando o que é a consciência, uma vez que a consciência é idêntica aos processos cerebrais. Não há

problemas quanto à forma como a mente está relacionada ao mundo físico; ela é meramente um aspecto ou subconjunto de eventos materiais. Esta abordagem da mente, uma segunda teoria do problema mente/corpo, conhecida como a teoria da identidade mente/cérebro, constitui o tópico do Capítulo 3.

No entanto, tanto o dualismo quanto a teoria da identidade mente/cérebro são igual e opositivamente erradas de acordo com uma terceira teoria, a do behaviorismo\* lógico ou analítico. A mente não é uma coisa fantasmagoricamente imaterial, mas tampouco o cérebro é material bruto. O modo mais frutífero de construir a mente é pensá-la como um padrão de comportamento real e possível exibido tanto por seres humanos quanto por animais. Os desempenhos da mente de uma pessoa não se encontram escondidos por detrás de seu comportamento, mas encontram sua corporificação dentro dela. Esta teoria, que se tornou influente a partir da década de 1930, por 25 anos ou mais, será examinada no Capítulo 4.

O Capítulo 5 nos traz uma quarta teoria, o funcionalismo, que se desenvolveu, de forma muito independente, a partir do behaviorismo analítico, por um lado, e de teorias computacionais da mente, por outro. O funcionalismo tem sido a tentativa mais recente e na moda para explicar a mente, embora haja agora sinais de que sua popularidade esteja em baixa. De acordo com os funcionalistas, a mente é estritamente nem mental, nem física, mas deve ser concebida de modo mais abstrato, como uma função ou programa, rodado no hardware do cérebro, que transforma entradas sensoriais em saídas comportamentais.

Por encontrar que nenhuma das teorias mencionadas é satisfatória, o monis-

mo não-redutivo compõe o assunto do Capítulo 6. O materialismo está correto ao rejeitar a mente concebida como uma alma não-física, mas está crucialmente deficiente em sua recusa em reconhecer a existência da consciência como um traço da realidade que não pode ser reduzido a processos puramente físicos. No entanto, há um sério problema no que diz respeito à capacidade do monismo não-redutivo para acomodar e explicar a causação psicofísica. Um retorno ao aristotelismo no Capítulo 8, que aborda as pessoas como unidades psicofísicas, parece mais frutífero, finalmente. Os estados mentais são mais bem concebidos como capacidades, cujos veículos são os processos neurológicos que os tornam possíveis, e que assim restauram seus poderes causativos.

Além das seis principais teorias da mente, às quais me restrinjo por questões de espaço e porque este é um livro introdutório, há algumas importantes questões que permanecem e requerem um tratamento extenso e separado, por si mesmas. A primeira destas, que compõe o Capítulo 7, diz respeito à questão de como a mente pode afetar o corpo, quando sua natureza parece ser tão diferente da natureza deste. Como pode alguma coisa aparentemente não-física, como um pensamento contrangedor, fazer-me enrubescer, por exemplo? Em uma tentativa de responder a este e a outros problemas a ele relacionados, eu me concentro amplamente em uma discussão sobre a teoria complexa do monismo anômalo de Donald Davidson.

No Capítulo 9, mudo de questões metafísicas sobre a natureza da mente para questões epistemológicas concernentes à sua cognoscibilidade. Temos um conhecimento especial sobre a existência e a natureza de nossos próprios estados mentais que nos falta no caso de outros? Podemos,

\*N. de T.: Optou-se por manter o aportuguesamento de *behaviorism*, nome pelo qual a escola é mais conhecida, do que por utilizar o termo "comportamentalismo", sua tradução.



de fato, alguma vez saber sobre a existência de estados mentais que não sejam os nossos próprios, ou estamos condenados a um eterno ceticismo sobre tais questões?

Finalmente, nos Capítulos 10 e 11, que são dedicados à questão de identidade pessoal, promovo uma ampla discussão sobre o que constitui a identidade de uma pessoa ao longo do tempo. O que faz com que uma pessoa agora, no tempo presente, seja uma e a mesma de um tempo passado, independentemente das mudanças consideráveis pelas quais ela passou durante esse meio tempo? O que, crucialmente, não pode ser mudado, para que não mude a identidade da pessoa? Será a alma, o corpo, ou uma parte específica do corpo, o cérebro? Por outro lado, será que nenhuma dessas respostas está correta? Será, quem sabe, a continuidade da memória e do caráter, independentemente atingido, que é crucial para que alguém seja uma e a mesma pessoa agora como ela foi então? Tentar esclarecer esta questão, que por si só já é de grande interesse e fascinante, é essencial para que se obtenha uma melhor compreensão da natureza do sujeito, ou eu, que tem experiências.

### **Abordando o problema mente/corpo**

Na seção anterior, delineei seis importantes teorias que disputam a solução do problema mente/corpo. Qual delas é correta, se é que alguma é, e quais são os critérios segundo os quais podemos verificar as adequações e as inadequações de teorias divergentes? Uma forma possível de começar a preparar esta tarefa é tentar descobrir que traços da mente e sua relação com o corpo uma teoria bem-sucedida deve ser capaz de explicar e acomodar. A ênfase neste primeiro estágio, salienta-se, recai, sobretudo, na descrição, e não na

avaliação. A avaliação detalhada de cada uma das teorias rivais da mente e do corpo ocorrerá em capítulos subsequentes, uma vez que os traços principais dos estados mentais já tenham sido caracterizados mais detalhadamente.

### **Características de estados mentais**

Os tipos de questões que devem ser tratadas nesta seção, bem como nas próximas, são os seguintes:

1. Quais são os traços distintivos dos estados mentais?
2. Todos os fenômenos mentais são semelhantes, ou há diferenças cruciais entre eles?
3. É possível detalhar os traços essenciais de vários tipos de estados mentais?

Para começar, necessitamos coletar um bom número de exemplares de diferentes tipos de estados mentais; em vista disso, antes de prosseguir na leitura, veja o Exercício 1.1.

### **EXERCÍCIO 1.1**

Faça intempestivamente uma lista de aproximadamente 15 itens que você classificaria como fenômenos ou estados mentais. Procure pensar uma grande variedade de exemplos – será muito proveitoso se você fizer este exercício com duas ou três outras pessoas.

Como você se saiu? Uma tentativa preliminar poderia ter produzido o seguinte:

Dores, comichões, sonhar, pensar, alucinar, ver, ouvir, ciúme, esperança, depressão, alegria, acreditar, querer, lembrar,

consciência, indignação, compreender, agir, intencionar, conceber, medo, desejar, lamentar, imaginar, cócegas, vibração, degustar, dores persistentes, cheirar, ver na imaginação, tocar, felicidade, miséria, inveja, raiva, ver pós-imagens, tentar, antecipar, pesar, desfrutar, formigamento, saber, raciocinar.

Vamos agora tentar uma tarefa um pouco mais difícil, como a proposta no Exercício 1.2.

### **EXERCÍCIO 1.2**

Percorra os exemplos de estados mentais dados anteriormente e verifique se pode encontrar modos de classificá-los em grupos diferentes. Por exemplo, alguns se referem a sensações, tais como dor e comichão, enquanto outros referem-se a emoções, e ainda outros são mais intelectuais, tais como pensar e compreender. Novamente, faça este exercício com duas ou três outras pessoas e depois compare seus resultados com os meus.

Classifiquei os estados mentais sob seis categorias:

*Sensações* – dores, dores persistentes, cócegas, vibração, formigamento.

*Cognições* – acreditar, saber, compreender, conceber, pensar, raciocinar.

*Emoções* – medo, ciúme, inveja, raiva, pesar, indignação, desfrutar.

*Percepções* – ver, ouvir, degustar, cheirar, tocar.

*Estados de quase-percepção* – sonhar, imaginar, ver na imaginação, alucinar, ver pós-imagens.

*Estados conativos* – agir, tentar, querer, intencionar, desejar.

Uma inspeção desta lista revela algumas importantes diferenças entre os itens que nela se encontram, como o restante desta seção aponta.

### **Localização corporal**

As sensações têm uma localização corporal mais ou menos definida, enquanto todos os demais estados não têm. Faz sentido dizer que você tem uma dor no seu lado ou um desconforto no seu ombro, mas não faz sentido dizer que sua crença está bem acima de sua orelha direita, seu ciúme no seu dedo mindinho ou sua intenção de visitar Paris no seu dedo do pé. No entanto, mesmo no caso de sensações, deve ser lembrado que a localização corporal delas não pode ser compreendida da mesma forma como falamos sobre a localização de objetos físicos. Se uma pessoa diz que tem uma dor no seu dedo mindinho e nós o abrimos para ver o seu interior, não encontraremos nada, nem mesmo um misterioso tipo de coisa chamada de dor. As dores têm uma localização somente no sentido de que elas estão onde quer que seja que a pessoa, que verdadeiramente sente a dor, refere que ela esteja.

A diferença entre falar a respeito da localização de sensações e a localização de objetos físicos pode ser ilustrada pelo Exercício 1.3, que você deveria fazer antes de prosseguir na leitura.

### **EXERCÍCIO 1.3**

Se minhas moedas estão na minha bolsa e minha bolsa está no meu bolso, segue-se que minhas moedas estão no meu bolso?

Em um sentido, minhas moedas não estão no meu bolso, isto é, não estão soltas no meu bolso, uma vez que se encontram dentro da bolsa. Mas claramente, se a bolsa está no espaço circunscrito pelos lados e fundo do bolso, e as moedas estão no espaço circunscrito pelo material que compõe a bolsa, então as moedas estão no meu bolso. A localização das moe-

das é uma localização literal no espaço, e alguém pode alcançar o bolso e retirar as moedas. O coração e os nervos, por exemplo, claramente têm localizações literais dentro do corpo e poderiam ser retirados dele.

Vejamos agora o Exercício 1.4.

#### **EXERCÍCIO 1.4**

Se minha dor é no meu pé, e meu pé está no meu sapato, segue-se que minha dor está no meu sapato? Este caso é muito mais complicado do que o do Exercício 1.3.

Em primeiro lugar, a minha dor não está no meu pé do mesmo modo que minhas moedas estão em minha bolsa. Eu posso tirar as moedas de dentro da minha bolsa, mas não posso chegar no meu pé e remover minha dor. Minha dor não é um objeto como uma moeda. A dor está no meu pé no sentido em que meu pé dói. Se eu digo que estou com dor e alguém então me pergunta onde é que dói, espontaneamente apontarei para meu pé. Mas isto não é como dar a localização literal no espaço de um objeto tal como uma moeda. Diz-se que a dor tem uma localização fenomenal – uma localização como indicada pelo sujeito que a experiencia. Pessoas que tiveram algum de seus membros amputados, frequentemente referem sentir dor e outras sensações no membro não-existente – isto é chamado de dor de membro-fantasma. Consequentemente, um amputado apontará para uma região de espaço vazio, e referirá que é lá que a dor está. Embora isto pareça estranho, faz sentido, uma vez que você lembre que especificar a localização de uma dor não é o mesmo que dizer onde está um objeto tal como uma moeda. O fato

de que a dor carece de uma localização espacial, diferentemente de objetos materiais, fornece mais uma razão para que não se conclua que uma dor existe no seu sapato, quando é seu pé que está sendo machucado no seu sapato. Finalmente, caso você insistisse que sua dor está no seu sapato, teríamos de lembrá-lo de que isto resulta em dizer que seu sapato está com dor. Isto é manifestamente absurdo, uma vez que sapatos, diferentemente de pessoas e animais, não são sujeitos que tenham experiências, sensações ou consciência.

#### **Sensações e ter consciência de algo\***

O caso esboçado no Exercício 1.4 pode ser usado para ilustrar outra diferença crucial entre sensações e objetos físicos. As moedas em seu bolso existem, quer você esteja consciente disso ou não. No entanto, seria autocontraditório dizer que você tinha uma dor no pé, mas que você não podia sentir nada desagradável acontecendo naquela parte do seu corpo. Dores e outras sensações existem somente quando você tem consciência delas. Ter consciência ou percepção delas faz parte integral da própria existência delas. Não é algo opcional ou um extra.<sup>3</sup>

#### **Estados de não-sensação e ter consciência de algo**

Os melhores exemplos de estados de não-sensação são fornecidos por estados cognitivos, que têm a ver com a compreensão e com o pensamento, e estados conativos (derivado do latim *conatus*, que significa esforço ou tentativa), que direta ou indiretamente dizem respeito a agir, querer tentar, querer e intencionar.

\*N. de T.: No original: *awareness*. Termo em inglês sem tradução adequada ao português. Seu campo semântico cobre os de ter consciência de alguma coisa, de perceber.

Em contraste com as sensações, é possível termos crenças, conhecimento e compreensão, bem como desejos e intenções, das quais não estamos conscientes em um dado momento. Assim sendo, tenho a intenção de visitar a República Tcheca neste verão, mas até que eu precisasse de um exemplo para ilustrar o ponto que estou levantando aqui, não tive a intenção de forma clara em minha mente. Do mesmo modo acontece com crenças e conhecimento. Há literalmente centenas de coisas nas quais acredito e que sei, mas, neste momento, não tenho consciência do que elas sejam. Para que tais crenças e conhecimentos possam ser atribuídos verdadeiramente a mim, ao menos deve ser possível que eu os possa trazer à minha mente se, e quando, assim for necessário. Isto é a que me refiro ao dizer que crenças e conhecimento são disposicionais, em vez de, como no caso das sensações, ocorrentes.

Uma sensação tal como a de dor ou cócegas existe se, e somente se, você está ocorrentemente consciente dela, isto é, continuamente consciente dela. Ao contrário, pode-se dizer que uma pessoa tenha uma crença ou intenção se, e somente se, há circunstâncias nas quais ela poderia trazê-la à mente, quando ela estivesse inclinada ou disposta a trazê-la à consciência, não que ela esteja realmente trazendo-a à mente. Dito de outra forma, dores e outras sensações existem somente quando você tem realmente consciência delas. No caso de crenças e intenções, bem como de outros estados de não-sensação, a imposição é mais fraca. Para que crenças e intenções sejam verdadeiramente atribuídas a uma pessoa, é suficiente que ela seja potencialmente objeto de consciência. Caso não haja quaisquer circunstâncias em que uma crença ou intenção possa ser trazida à mente e expressa, então é altamente improvável que tal crença ou intenção exista.<sup>4</sup>

Emoções e estados de espírito são mais difíceis de serem tratados. Se você está brabo em um dado momento, segue-se que você deve estar consciente de sua brabeza e ciente de seus sentimentos. Por outro lado, parece possível que se esteja brabo com alguma pessoa, sem que constantemente se tenha consciência disso. Você poderia permanecer brabo com alguém durante dias, e sua brabeza perdurar até durante seus períodos de sono. Isto se deve ao fato de as emoções, à semelhança de crenças e intenções, serem de caráter disposicional.

### **Qualia e estados mentais**

No caso de sensações, há algo que se assemelha a experimentá-las. Pense na sensação de comichão de um cobertor de lã bruta, a sensação de picada produzida por uma agulha, o frescor calmante de uma loção para queimadura solar, a sensação de formigamento de alfinetes e agulhas, ou o latejar de um dente cariado. Os filósofos falam da fenomenologia das sensações, isto é, a forma como as coisas parecem ao sujeito que as experienciam, cada tipo de sensação sentida de um modo qualitativamente distinto. Este segundo modo de caracterizar as sensações levou os filósofos a inventar uma expressão técnica especial e a descrever as sensações como possuidoras de "qualia" ou "sensações em estado bruto".

Em contraste com as sensações, parece não haver fenomenologia distintiva alguma associada a crenças ou outros estados cognitivos, tais como a compreensão e o pensamento. A crença de que o Royal Festival Hall está na margem direita do rio Tâmisa não tem qualquer aspecto ou sensação associado com ela. Certamente, quando afirmo esta crença, memórias associadas às minhas frequentes visitas a ele podem permanecer em minha mente, mas

estas são puramente incidentais à crença e não têm papel algum na constituição dela. Mesmo se nenhuma lembrança ou associação é invocada quando o Festival Hall é mencionado, pode-se, no entanto, ainda dizer que eu acredito que ele está na margem direita do Tâmisa.

Diferentemente de crenças, e mais parecidas com sensações, as emoções são comumente acompanhadas de formas de sentir, isto é, elas têm uma fenomenologia distintiva, que lhes é própria. Assim, nós falamos de sofrer ataques de ciúmes, de ficar roxo de raiva, de ficar agitado por uma roedora preocupação, de uma sensação horrorosa na boca do estômago e assim por diante. Há outros fenômenos mentais que se parecem com as emoções, por terem uma fenomenologia. Por exemplo, sofremos os ataques de lamento ou de remorso, compartilhamos a excitação da antecipação, gratificamo-nos na nostalgia de lembranças amargas ou doces e sentimos a desolação de esperanças que não deram em nada.

Alguns estados conativos também parecem ter uma fenomenologia associada a eles, como quando falamos do fogo do desejo, do esforço envolvido na tentativa de lograr êxito em uma tarefa difícil ou perigosa, ou quando fervorosamente desejamos que algo venha a se realizar. Outras, tais como as intenções, estão mais próximas das crenças e, para adotar uma expressão usada originalmente por Ludwig Wittgenstein (1889-1951) a respeito de ações, parecem ter "nenhum volume de experiência", isto é, não há conteúdo qualitativo algum que permita caracterizar intenções.

### **Sobre a "sobredade"**\*

Examinando os vários tipos de estados mentais descritos anteriormente,

uma distinção de vital importância que se apresenta é aquela entre estados mentais, que parecem ser "sobre" outros estados de coisas, ou a eles dirigidos, e aqueles que não são. Sensações não são a respeito de outros estados de coisas, reais ou imaginários, ou dirigidas a eles, que sejam externas a si mesmas. Em outras palavras, sensações não representam outros estados de coisas possíveis. Ficaríamos atônitos se nos perguntassem a respeito de que são as nossas dores ou cócegas, em oposição ao que as estavam causando. Sensações, estaríamos inclinados a dizer, simplesmente são. Elas não são a respeito de alguma coisa, assim como a mesa e a cadeira nas quais estou trabalhando agora não representam ou retratam outros estados possíveis de coisas.

Contraste isto com a crença. Você não pode simplesmente acreditar e ponto final. Deve haver algum conteúdo para a sua crença, ela diz respeito a alguma coisa, independentemente de haver realmente no mundo algo que corresponda àquele conteúdo. Você pode acreditar que a Lua é feita de queijo roquefort, por exemplo. Ela certamente não é; no entanto, "que a Lua é feita de queijo roquefort", um estado de coisas não-existente, é a que a sua crença diz respeito, o que a crença representa para você como é a Lua. Desejos e intenções, à semelhança da crença, também possuem uma "sobredade". Você não pode simplesmente querer ou intencionar o que é que você quer ou objetiva fazer. O que torna uma crença verdadeira é se o mundo realmente é como a crença o representa ser. O que realiza uma intenção é a concretização intencional, pela pessoa que teve a intenção, do estado de coisas no mundo de acordo com a especificação dada. O filósofo John Searle falou em relação a isto de "direções de adequação" de

\*N. de T: No original: *aboutness*, uma forma substantivada de *about*, aqui traduzido por "sobredade"; em outros contextos e tradutores, por *intencionalidade*.

estados que comportam uma “sobredade”. No caso de crenças, a direção do ajuste vai do mundo à mente; como o mundo é dito, se a crença é verdadeira ou não. No que diz respeito a intenções, a direção de ajuste vai da mente para o mundo: a intenção, supõe-se, faz com que o mundo se adapte a seus requisitos.

Emoções, à semelhança de crenças e intenções, também devem ser direcionadas a um conteúdo. Se, por exemplo, você está brabo, deve haver alguém com quem você está brabo, ou algo a respeito do qual você está brabo. Mas emoções são mais complexas do que crenças, porque elas pressupõem crenças e não podem ser reduzidas a estas. Uma crença é um componente de uma emoção, não a própria emoção. Para ilustrar este ponto, suponha que você está brabo com alguém que traiu a sua confiança. Segue-se, portanto, que você deve acreditar que a pessoa em pauta traiu a sua confiança. Na ausência desta crença, ou no caso de você descobrir que esta crença é falsa, você não ficará brabo. Os filósofos quiseram fazer a distinção entre objeto de uma emoção – sobre o que é a emoção, ou a que ela se direciona – e a razão pela qual a emoção é sentida, esta razão sendo suprida pela crença. A filósofa Elizabeth Anscombe (1921-2001) dá o exemplo de uma criança que estava assustada com um pedaço de cetim vermelho deixado sobre uma escada. Por que deveria ela estar aterrorizada por um pedaço de tecido? Resposta: ela acreditava que lhe houvessem dito que aquilo era um pedaço do Diabo. Uma vez que a criança seja libertada desta crença, então seu medo, se for racional, se evaporará.

Os filósofos falam frequentemente usando um termo técnico especial, não a “sobredade” dos estados mentais, tais como crenças, desejos, emoções e intenções, mas da “intencionalidade” deles.

Assim, tais estados são conhecidos como estados intencionais, em contraste com dores e outras sensações que são de caráter não-intencional. O conteúdo de tais estados, sobre o que eles são, é chamado de conteúdo intencional. Uma explicação mais completa da intencionalidade e de como este termo veio a ser criado por filósofos medievais será dada na próxima seção.

Ações também apresentam intencionalidade. Embora haja coisas que fazemos sem uma razão em particular, por hábito ou chateação, como manusear nervosamente um clipe de papel ou nosso cabelo enquanto estamos falando, via de regra as ações são tomadas para a realização de propósitos. Neste sentido, são dirigidas a um fim, do mesmo modo que as intenções. Mas as ações são físicas ou mentais? Certamente há ações que são puramente mentais, no sentido de que não envolvem movimento corporal algum. Fazer cálculos aritméticos mentais, tirar conclusões a partir de premissas e detectar pressuposições de argumentos são todas atividades que podem ocorrer internamente sem qualquer movimento físico. Outras atividades, como levantar o braço, necessariamente envolvem movimentos corporais. Tais atividades não são, no entanto, idênticas aos próprios movimentos. Se eu levanto o meu braço, segue-se que meu braço se levanta, mas se meu braço se levanta, não se segue que eu o levantei. Meu braço poderia ter se levantado por razões não conectadas à minha ação pessoal – alguma outra pessoa o poderia ter levantado, ou eu poderia ter sofrido um espasmo. Parece, então, que as ações são de natureza tanto física quanto mental, distintamente dos movimentos corporais, que não têm aspecto mental algum.

Percepções e quase-percepções também exibem “sobredade”. Para mim, um

estado quase-perceptivo quer dizer um estado que parece à pessoa que está realmente ouvindo ou vendo alguma coisa, enquanto de fato ela está somente sonhando, imaginando ou alucinando. Seus sonhos, imaginações e alucinações têm um conteúdo mental, mas, à semelhança dos casos de falsas crenças, tal conteúdo não representa um estado de coisas realmente existente no mundo. Por outro lado, no que diz respeito a estados de percepção genuínos, tais como ver e ouvir, por exemplo, o que é visto ou ouvido deve realmente existir. Esta é a consequência das condições que estados de percepção devem satisfazer para que possam ser qualificados como estados de percepção. Para que se possa realmente dizer que eu vi alguma coisa, então não só tal coisa deve existir para ser vista, como também deve desempenhar um papel apropriado em suscitar minha percepção dela. Caso esta condição não seja satisfeita, então teríamos de descrever-me como alguém que parece estar vendo a coisa em questão – talvez eu a esteja imaginando ou sonhando.

Seria tentador identificar o conteúdo intencional da experiência de ver um objeto com o próprio objeto, mas isto seria um engano. Isso porque uma distinção pode ser estabelecida entre como um objeto aparece para nós – o modo, poderíamos dizer, como ele apresenta a si próprio para nós – e como ele realmente é. Alguns exemplos podem ajudar a esclarecer este ponto. O Sol parece ter aproximadamente o tamanho de uma moeda de cinco centavos, mas sabemos que ele é muito maior do que a Terra. A grama quando vista à distância parece ser azul, mas, de perto, vemos que ela é verde. Portanto, há quadros ambíguos que podem ser visto de modos diferentes. O desenho mostrado na Figura 1.1 pode primeiramente ser visto como um pato e depois como um coelho, ou talvez

para algumas pessoas somente algumas linhas sobre o papel. Embora haja um sentido no qual tanto você quanto eu vemos somente um e o mesmo quadro, o conteúdo intencional de nossas respectivas percepções, o modo como o quadro apresenta a si próprio para nós, é diferente para cada um de nós. Você o vê como sendo um pato, eu como um coelho. Em outras palavras, percebemos diferentes aspectos, o quadro como tal, naturalmente, permanecendo imutável todo o tempo. O exemplo do pato/coelho ajuda a explicar o que se quer dizer ao descrever a percepção visual como tendo um caráter de aspecto, e esclarece por que o conteúdo intencional da percepção não pode simplesmente ser identificado com o objeto externo percebido.

O exemplo do pato/coelho ilustra a situação em que temos um e o mesmo quadro, que apresenta dois aspectos diferentes. No entanto, não é difícil pensar a possibilidade oposta, em que duas coisas diferentes apresentem o mesmo aspecto. Rosie e Alice são gêmeas idênticas, impossíveis de distinguir. Para se divertirem, as gêmeas fazem uma brincadeira em que Rosie me diz ser Alice e Alice me diz que é Rosie. Assim, frequentemente parece que eu estou falando com Rosie, quando de fato estou falando com Alice, e vice-versa. O conteúdo intencional de minha percepção será o mesmo em ambos os casos, embora na realidade eu esteja vendo e falando com duas pessoas diferentes.

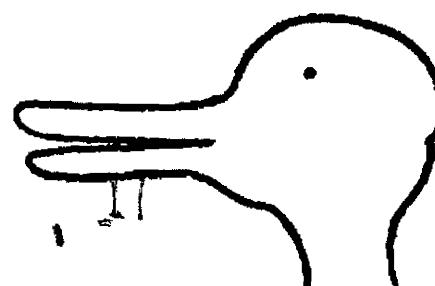


Figura 1.1

A aspectualidade de estados de percepção não se restringe exclusivamente a casos de “ver como”, ou, fazendo uma colocação mais geral, a casos de “ter experiência como”. Há também “ver que”, para restringir a discussão a somente um dos sentidos. Assim, após um longo tempo de convivência, finalmente logrei distinguir Rosie e Alice, e não posso mais me deixar enganar, uma vez que agora posso ver que é Rosie e não Alice que está diante de mim.

Há sempre a tentação de se pensar o conteúdo intencional de uma percepção como se fosse um objeto mental que se coloca entre nós e o objeto visto, como se o modo como um objeto aparece ou parece fossem como uma membrana que pode ser retirada do objeto e contemplada independentemente dele. Caso este movimento seja feito, então o conteúdo perceptual é então dito representar o objeto que está sendo percebido. O problema com este modo de perceber o conteúdo perceptual é que o conteúdo se torna um substituto, uma substituição do próprio objeto, e esse substituto então se coloca entre nós e o objeto. Ficamos então com o problema de explicar como, se temos conhecimento somente do substituto que representa o objeto, jamais viremos a saber que realmente haja em algum lugar o objeto que faça com que a representação ou substituto exista em primeiro lugar. O antídoto para isto é dar-se conta de que o modo como um objeto parece ou aparece não consiste em uma entidade misteriosa chamada uma “aparência”, que nos tira de contato com o próprio objeto. O conteúdo intencional de uma percepção é simplesmente o próprio objeto, que parece ou aparece de certo modo para nós. Por esta razão, os filósofos têm recomendado que o conteúdo intencional das percepções não deve ser pensado como uma representação do mundo para nós, mas, sim, como uma apresenta-

ção dele de uma certa maneira. Falar de representar estimula a crença enganosa de que na percepção estamos realmente apenas cientes de entidades chamadas de representações, que se colocam entre nós e o objeto que está sendo percebido, em vez do próprio objeto.

### **Intencionalidade**

De acordo com o que foi prometido na seção “Sobre a ‘sobredade’”, agora direi um pouco mais sobre o conceito de “sobredade”, ou intencionalidade, que era um traço da maioria dos vários tipos de estados mentais examinados.

O adjetivo “intencional” e o substantivo abstrato “intencionalidade”, derivado daquele, foram cunhados pelos filósofos medievais ou escolásticos. A expressão se deriva do latim *intendere* que significa “visar”, no sentido de esticar o arco a fim de disparar uma flecha em um alvo (*intendere ad arcum*). Os termos “intencional” e “intencionalmente”, comumente utilizados para descrever ações, derivam-se da mesma origem, e retêm o sentido de um objetivo ou alvo ao qual a ação é visada. No entanto, “intencionalidade” é um termo filosófico especial que é importante não seja confundido com o “intencionalmente” de todos os dias. Após haver sido um tanto negligenciado depois da Idade Média, o conceito de intencionalidade foi revivido por Franz Brentano (1838-1917), que escreveu:

Todo fenômeno mental é caracterizado pelo que os escolásticos da Idade Média chamavam o Intencional (e também mental) inextensivo de um objeto, e que chamaríamos, embora em termos não inteiramente isento de ambiguidade, a referência para um conteúdo, uma direção sobre um objeto (pelo qual não devemos compreender uma realidade neste caso) ou objetividade imanente. Todo fenômeno

mental inclui algo como objeto dentro de si mesmo, embora nem todos eles façam isso da mesma maneira. Na apresentação, algo é apresentado, no julgamento, algo é afirmado ou negado, no amor, amado, no ódio, odiado, no desejo, desejado, e assim por diante. Esta inexistência intencional é característica exclusivamente dos fenômenos mentais.<sup>6</sup>

Esta passagem carece de um breve comentário. É pelo termo ligeiramente obscuro “inexistência intencional” que Brentano chama o que agora é chamado de conteúdo representativo de um estado intencional, o estado de coisas, real ou imaginário, que o estado representa. No entanto, como Brentano assinala mais adiante, antecipando Searle (ver seção “Sobre a ‘sobredade’”), nem todos os conteúdos são apresentados da mesma forma. Uma crença, por exemplo, afirma que o mundo é como a crença descreve que ele é. Desejos e intenções, por outro lado, especificam como o mundo deveria se tornar para que se conformasse ao

conteúdo do desejo ou intenção em pau-  
ta. Desde que o filósofo Bertrand Russell (1872-1970) inventou o termo, tornou-se comum chamar o modo de apresentação de um conteúdo intencional de uma “atitu-  
de”, uma instância mental, por assim dizer, tida em direção a um dado conteú-  
do por um sujeito e caracterizada por um verbo psicológico. O próprio conteúdo é usualmente especificado por uma ora-  
ção iniciada por “que”, que introduz uma proposição ou afirmação, ou, alternativa-  
mente, por uma frase começando pelo in-  
finitivo de um verbo. Assim, os estados inten-  
cionais são também conhecidos como as atitudes proposicionais. O Quadro 1.1 nos auxilia a tornar isto mais claro.

Brentano pode ser culpado de exage-  
ro, no entanto, quando ele afirmava que todo fenômeno mental é caracterizado pela referência a um conteúdo, uma direção so-  
bre um objeto, uma vez que, como salien-  
tei na seção anterior, as sensações, diferen-  
temente das crenças e das intenções, não  
representam estados de coisas externas a

**Quadro 1.1** Estados intencionais (atitudes proposicionais)

| Atitude  | Conteúdo proposicional   |
|--|--|
| <i>Oração principal especificando o tipo de estado intencional em questão por meio de um verbo psicológico que descreve a atitude adotada em direção ao conteúdo</i> | <i>Oração-que encaixada ou infinitivo de um verbo especificando o conteúdo proposicional (representacional) do estado intencional – a respeito de estado intencional</i> |
| Eu creio   | que alguém em algum lugar se importa comigo  |
| Lois Lane imagina  | que está beijando o Super-Homem  |
| Ele quer   | visitar as Malvinas  |
| Eles têm a intenção de   | viajar para o exterior nas suas férias   |
| Ele duvida   | que Deus existe  |

si mesmas. No entanto, as outras afirmações de Brentano, particularmente a de que a intencionalidade se aplica exclusivamente a fenômenos mentais, podem ser postas em questão, porque a linguagem também parece ser intencional. Este livro, por exemplo, é a respeito da natureza da mente. As ficções de Terry Pratchett são a respeito de um Mundo Disco místico não existente, que se move pelo espaço carregado às costas de um gigante elefante. De forma semelhante, programas de computador e impressos, filmes, fotografias, mapas, diagramas, pinturas, modelos, esculturas e esfregações sobre cobre também representam estados de coisas, reais ou imaginários, distintos de si próprios.

Antes de prosseguir na sua leitura, veja o Exercício 1.5.

### **EXERCÍCIO 1.5**

Você pode pensar em uma objeção à afirmação de que coisas que não são estados mentais, tais como palavras, exigem intencionalidade? Trabalhando sozinho ou com alguém mais, escreva sua objeção.

A objeção a que você pode chegar é a de que a linguagem, considerada puramente em si mesma, não possui intencionalidade original ou intrínseca. A intencionalidade intrínseca é um traço exclusivo das mentes de humanos (e de animais). As marcas ou os ruídos que vão na direção de compor uma linguagem têm uma intencionalidade meramente derivativa ou “como se”, que é atribuída a eles por usuários da linguagem dotados de mentes. As palavras têm significado e podem ser a respeito de coisas somente em relação a pessoas para as quais elas têm um significado e que podem interpretá-los como sendo a respeito de vá-

rios estados de coisas. Imagine que toda a vida consciente no universo cessasse repentina e permanentemente, mas que todos os livros permanecessem intactos. Em si mesmos, intrinsecamente, nenhum desses itens possuiria significado ou representaria possíveis estados de coisas fora de si próprios. As marcas de tinta que compõem as frases escritas nos livros ainda existiriam, mas com o desaparecimento de alguém para ler e compreender tais frases todo o significado teria evaporado. Tudo o que existiria basicamente seriam pedaços de papel com marcas pretas sobre eles. Analogamente, os documentos impressos por um computador não significam nada para o próprio computador, uma vez que computadores, carecendo de estados mentais, não possuem intencionalidade intrínseca. Esta é a posição adotada pelo pensador norte-americano John Searle. No entanto, como veremos mais tarde, há filósofos que vigorosamente contestariam estas afirmações. Eles veriam a afirmação de que computadores não podem possuir uma intencionalidade genuína como nada mais do que um preconceito, que meramente exclui a possibilidade de que uma intencionalidade genuína possa ser encontrada em algum outro lugar à exceção das mentes humanas e de animais.

A tentativa de explicar como a intencionalidade é possível tornou-se uma importante indústria filosófica desde os anos 1970, aproximadamente. Particularmente, o objetivo de muitos filósofos que escrevem sobre esse assunto tem sido o de demonstrar como a intencionalidade pode ser naturalizada. Embora não haja definição alguma de naturalismo que seja consensual a todos os filósofos, em termos gerais, a ideia é de que, ao naturalizar a intencionalidade, não necessitamos lançar mão de estados fantasmagóricos não-físicos, que são de natureza supranatural, isto

é, fora ou acima do mundo físico ordinário, para explicar como os estados mentais podem ter um caráter intencional. Servindo-se de teorias e conceitos tirados das ciências, a esperança é de que se mostrará que a intencionalidade é somente um outro traço do mundo natural, como a digestão ou a secreção de bile. No entanto, embora venhamos a abordar algumas questões referentes à intencionalidade no nosso exame de teorias da relação mente/corpo, o seu exame completo está além do escopo deste livro.

### **As perspectivas de primeira e terceira pessoas e a alegada privacidade do mental**

Quando eu descrevo os traços físicos de uma outra pessoa, ou de fato os meus próprios, faço isso com base em uma experiência sensorial, ou seja, por meio da visão, do tato, do olfato e assim por diante. Do mesmo modo, quando alguém descreve como eu sou fisicamente, chega à sua caracterização fundamentalmente pelo mesmo processo. Nenhum de nós dispõe de algum meio especial ou, em princípio, privilegiado de descobrir os fatos físicos a respeito de outras pessoas e de nós próprios, que outros não tenham. O mesmo se aplica caso eu deseje descobrir o funcionamento interno de meu próprio corpo, ou o de alguma outra pessoa: Se quisesse descobrir o que estava acontecendo no meu cérebro, por exemplo, eu não teria nenhum outro recurso senão o emprego, ou, o que é mais provável, encontrar alguém que utilize em meu lugar, essencialmente das mesmas técnicas que eu teria de usar para descobrir o que se estava passando em seu cérebro. O método mais direto, certamente, seria abrir o crânio e observar diretamente o cérebro vivo. Eu poderia até solicitar que isto fosse feito no

meu próprio caso – pacientes sob anestesia local têm visto seus próprios cérebros, durante cirurgias, refletido em espelhos convenientemente colocados ou em uma tela de vídeo. De modo menos radical, eventos em meu próprio cérebro, assim como no de outros, podem ser testados por meio de escâner ou, alternativamente, de eletroencefalogramas, que detectam e medem os distintos ritmos da atividade elétrica do cérebro. Podemos resumir esta posição dizendo que o conhecimento do funcionamento de nossos corpos, e a natureza física em geral, quem quer que sejamos, é feito a partir da perspectiva de uma terceira pessoa, não havendo nenhuma outra perspectiva à nossa disposição em tais casos.

Os casos de cérebros e mentes parecem ser muito distintos. Veja o Exercício 1.6 e procure encontrar as possíveis razões.

### **EXERCÍCIO 1.6**

O que é verdade em relação a cérebros é igualmente verdade quanto a mentes? Como descubro o que se está passando em minha mente, em oposição ao que está ocorrendo na sua? Há algum obstáculo para descobrir o que você está pensando e sentindo, comparado com a minha capacidade de dizer o que eu mesmo estou pensando e sentindo? Pense a respeito disso por alguns instantes e, se possível, discuta com alguma outra pessoa.

Em relação ao meu próprio caso, da forma mais usual e típica, eu não necessito recorrer a um processo de me observar na televisão ou no espelho para ser capaz de dizer, com precisão e veracidade, o que estou pensando ou sentindo. De fato, eu absolutamente não tenho de recorrer a qualquer procedimento de descoberta

dessas coisas. Parece que eu simplesmente tenho ciência ou consciência de meus pensamentos e sensações. No entanto, em relação a uma outra pessoa, digamos você, o único meio de eu saber quais são seus pensamentos e sentimentos é baseado na sua conduta, no que você diz ou faz, uma vez que eu pareço não dispor de qualquer outro meio mais direto de ter acesso a seus pensamentos e sentimentos. Além disso, caso você decida manter seus pensamentos só para você, talvez eu jamais possa saber quais são eles. Indo um pouco além, como posso me assegurar de que seu comportamento não é somente uma fachada, que encerra um gélido vazio mental? Você é, talvez, um zumbi sem mente, ou uma marionete sem consciência controlada por Marte (tomando emprestado um exemplo do filósofo Christopher Peacocke).

Esta linha de pensamento estimula fortemente a crença de que os estados mentais são radicalmente diferentes dos estados físicos. Os estados mentais são privativos do indivíduo, enquanto estados físicos, ao contrário, são públicos. Além do mais, esta privacidade não é igual à privacidade dos conteúdos do meu diário, ou o código no qual Samuel Pepys escondeu seus pensamentos concupiscentes. O código de Pepys foi subsequentemente quebrado sem dificuldade, e com a mesma facilidade alguém poderia roubar o meu diário e ler seus enfadonhos conteúdos. A privacidade da mente tem sido caracterizada como uma privacidade lógica, que é logicamente inviolável, de forma que ninguém, em princípio, jamais possa ter um pequeno acesso direto ao conteúdo da mente de alguma outra pessoa.

Assim, enquanto há uma simetria no que concerne ao meu e ao seu acesso a estados físicos, no sentido de que nenhum de nós tem um acesso privilegiado àqueles estados que sejam negados ao outro, este não é o caso quanto aos estados mentais.

Eu tenho um acesso aos meus próprios estados mentais, que lhe é negado, assim como você tem acesso aos seus próprios estados mentais, o que me é negado. Em outras palavras, há uma assimetria no que concerne ao acesso do que cada pessoa desfruta de seus próprios estados mentais, em contraste com o acesso que ele tem dos estados mentais de outras pessoas.

Isto é retratado por Ludwig Wittgenstein na sua metáfora do besouro-numa-caixa. Devo ser visto como alguém que tem uma caixa em que somente eu, em princípio, posso olhar – e todas as demais pessoas têm uma caixa semelhante. Todos nós olhamos dentro de nossas próprias caixas e reportamos que ela contém um besouro, mas não podemos mostrar o que está na caixa para os outros. É possível que a coisa que cada caixa contém seja muito diferente das outras, mas nunca descobriremos, uma vez que não podemos comparar os besouros diretamente. Se de algum modo eu pudesse somente olhar dentro de sua caixa, poderia fazer a alarmante descoberta de que o que você chama de besouro eu refiro como sendo uma aranha. O conteúdo da experiência de cada pessoa, a fenomenologia ou qualia que engloba seus estados mentais, poderia diferir drasticamente de uma pessoa a outra, podemos dizer. Não temos, em vista da própria natureza do caso, modo algum de descobri-lo. No limite máximo, a caixa pode até estar vazia, e não haver qualia algum, mas meramente um deserto mental. Além disso, uma vez que o significado das palavras que usamos é determinado pelos objetos aos quais elas se aplicam, uma consequência ainda mais alarmante se segue, qual seja, que cada um de nós poderia estar falando uma língua privada, os significados de cujos termos nunca poderiam em princípio ser profundamente apreendidos por outros. Que tal língua possa existir, uma língua logicamente

privada, como tem sido chamada, e que os conteúdos da experiência possam ser radicalmente privados segundo o modo descrito, tem sido ferrenhamente contestado por Wittgenstein e aqueles filósofos por ele influenciados. Voltarei a examinar estas questões mais detalhadamente no Capítulo 9.

### **Translucidez, incorrigibilidade e autoridade máxima**

Historicamente, a doutrina da privacidade e do acesso privilegiado tem sido associada a Descartes, assim como àqueles por ele influenciados a esse respeito, tais como Locke e Hume. Uma extensão natural da doutrina é a tese da translucidez, a afirmação de que uma pessoa não pode ser ignorante de seus próprios estados mentais, ou errada quanto aos tipos de estados mentais que são. O conhecimento que uma pessoa tem de seus próprios estados mentais, em contraposição ao seu conhecimento de outros assuntos, é não só completo, mas incorrigível ou infalível, isto é, além da possibilidade do erro ou da correção. Segundo a perspicua colocação de John Heil, a afirmação é a de que "eu conheço infalível e completamente todo e somente o conteúdo de minha mente. Onde p é uma proposição atribuindo um pensamento atual a mim próprio, é necessariamente verdadeiro tanto que, se eu acredito que p, então p é verdadeiro, e que se p é verdadeiro, eu acredito que p".<sup>7</sup> Hume, um cartesiano sob este aspecto, assume a seguinte atitude: "Uma vez que é por meio da consciência que conhecemos todas as ações da mente, elas devem necessariamente aparecer em todos os particulares que elas são e ser o que elas parecem". Levando a tese da translucidez ainda mais adiante, tem sido pensado com frequência que os conteúdos da mente de

uma pessoa são, de algum modo, autoevidentes. Eles positivamente clamam ser notados, e as metáforas utilizadas para caracterizar esta autopropaganda incluem pensar os itens mentais como jóias brilhando em um estojo, ou assemelhando-se a uma água marinha fosforescente.

Há, no entanto, muitos desafios que podem ser feitos à afirmação de que uma pessoa conhece infalivelmente seus próprios estados mentais. Considere inicialmente o caso da dor e de outras sensações. Se eu estou com dor, segue-se que eu sei que eu estou com dor? Aqui devemos ter cuidado. É verdade que sentir dor é simplesmente estar ciente de uma sensação dolorosa. A noção de uma dor não sentida, uma dor da qual não se tenha percepção ou consciência, é certamente, em termos, uma contradição. Neste sentido limitado, a ignorância de suas próprias sensações não é possível, porque, como vimos na seção sobre estados de não-sensação, a percepção das sensações é parte integral da existência delas.

No entanto, estar com dor é uma coisa, mas ter a percepção de que se está com dor é algo inteiramente diferente. A percepção de que se está com dor envolve um julgamento, enquanto simplesmente ter uma dor não envolve. Julgar que você está com dor significa que você está atribuindo uma dor para si mesmo, o que por sua vez pareceria envolver não somente a posse do conceito de dor, mas, em acréscimo, o conceito de si mesmo como um sujeito de experiência. Fazer julgamentos fazendo uso desses conceitos é uma atividade essencialmente linguística e, portanto, impossível para aqueles que não fazem uso da linguagem, tais como cães e gatos. Os animais podem sentir dor, mas disso não se segue que eles saibam que estão com dor, isto é, que eles possam julgar que estão com dor. Fazer um julgamento de que se está com dor envolve um passo além

do simples fato de experienciar a dor. Para aqueles que não podem dar tal passo, o conhecimento proposicional de seus próprios estados mentais – o tipo de conhecimento que figura na atribuição de estados mentais para si próprio – não é possível.<sup>9</sup>

### Avaliando as doutrinas do acesso privilegiado e da translucidez

#### EXERCÍCIO 1.7

Quão convincente você acha a afirmação de que uma pessoa tem acesso privilegiado a seus próprios estados mentais, e em que medida você concorda que o conhecimento que uma pessoa tem de seus próprios estados mentais seja infalível?

Parece que, no caso de ao menos alguns dos meus estados mentais, eu os posso atribuir a mim mesmo sem precisar fazer observações ou ter de me engajar em qualquer processo de tirar conclusões a partir de estados de coisas mais básicos. Dores e outras sensações propiciam alguns dos exemplos mais claros. Posso referir que estou com dor sem observação ou inferência. Outras pessoas, ao contrário, podem saber que eu estou com dor, se é que alguma vez poderão, somente baseados no que eu disser ou fizer. Eu também pareceria ser a autoridade máxima sobre como são as minhas sensações, ou se estou genuinamente as experiencingando. Desde que eu saiba o significado da palavra “dor”, parece que não posso estar enganado em dizer, sinceramente, que a perfuração que o dentista está fazendo está me machucando. Em um caso como este, sou certamente a autoridade máxi-

ma, e é difícil ver como terceiras pessoas poderiam estar em situação de negar ou corrigir, com sucesso, minha sincera afirmação de que estou com dor. O mesmo usualmente se aplica ao meu conhecimento de meus estados mentais desprovidos de sensações, tais como crenças, desejos e intenções. Normalmente eu posso confessar, sem medo de contradição ou correção, o que eu acredito ou pretendo fazer.

No entanto, quando se trata de estados intencionais, algumas vezes é bem menos do que certo que eu tenha melhor conhecimento do que outros sobre o que estou pensando e sentindo. Enquanto eu possa ser a autoridade máxima sobre se o sapato está apertando ou não, eu claramente não sou a autoridade máxima para dizer se comprehendo, por exemplo, algum complexo teorema de matemática. Caso eu fosse dizer que sei, com base no meu escrutínio de minha paisagem mental interior, que eu realmente comprehendo o teorema, ririam de mim, com justa razão, até o desprezo. O que mostrará se eu comprehendo o teorema ou não será revelado por coisas tais como minha capacidade de explicá-lo com minhas próprias palavras a alguma outra pessoa, de tirar corolários a partir dele, ou apontar outros teoremas que o contradigam. Novamente, eu não sou a melhor pessoa para fazer uma avaliação de meu próprio estado emocional. São recursos da ficção romântica, por exemplo, Mills e Boon, que as últimas pessoas a se darem conta de que estavam apaixonados um pelo outro são o herói e a heroína. Uma incapacidade de se conhecer a sua própria mente está bem ilustrada no *À la recherche du temps perdu*<sup>\*</sup>, de Proust. Bloch, um visitante na casa dos avós do narrador, ouve a avó, a quem ele nunca havia encontrado, reclamar que não estava se sentindo muito bem. Bloch ex-

\*N. de T.: Em português: *Em busca do tempo perdido*.

plode em lágrimas em consequência disso, o que levou a avó mais tarde a comentar que ele deve ser louco para exibir tal paroxismo de pesar a respeito de uma questão tão trivial. O ponto é que, embora Bloch se imagine sincero quanto às emoções que sente, ou tenha convencido a si mesmo de que é sincero, a profundidade de seus sentimentos e sua reação são desproporcionais ao seu objeto. Posteriormente, os pais do narrador concluíram que, embora Bloch fosse sincero no sentido de não estar fingindo sentir uma emoção quando na realidade nada sentia, no entanto,

nossas primeiras emoções impulsivas não têm mais do que uma pequena influência sobre nossas ações posteriores e sobre a conduta de nossas vidas; e este respeito pelas obrigações morais, lealdade aos amigos, obediência à regra da vida, têm um embasamento mais certo em hábitos solidamente formados e seguidos cegamente do que nestes transportes momentâneos, ardentes, mas estéreis.<sup>10</sup>

Outras variedades de autoengano também fornecem ricas instâncias de como nós podemos deliberadamente falhar na compreensão de nós mesmos. No *Lord Jim*, de Joseph Conrad, o epônimo herói consegue convencer-se de que a razão que o levou a não salvar os dois homens lançados ao mar foi ele ter sido pego de surpresa pela tempestade, em vez de ele ter ficado paralisado de medo ou covardia. Mais tarde, ele evita ter de confrontar seus ciúmes do rapaz que procedeu ao salvamento, etiquetando o feito deste como “uma lamentável mostra de vaidade”. E ele até consegue persuadir-se de que, ao não ir fazer o salvamento, ele “ampliou seu conhecimento mais do que aqueles que foram efetuar o trabalho” e obteve algo maior do que os salvadores, embora, ao mesmo tempo, em

seu coração, ele saiba perfeitamente bem que não é assim.<sup>11</sup>

John Heil, em *The Nature of True Minds\**, menciona ainda outras intrigantes possibilidades.<sup>12</sup> Poderia sustentar a crença de segunda ordem, a crença de que eu acredito em alguma proposição, p. Mas esta crença de segunda ordem poderia estar enganada em ao menos dois modos. Primeiramente, eu poderia não acreditar em p de forma alguma: em outras palavras, creio falsamente que mantendo a crença de que p. Em segundo lugar, enquanto adotei uma atitude proposicional em relação à p, poderia estar enganado quanto ao que seja esta atitude. Talvez, em vez disso, eu não acredite realmente que p, mas receie que p, ou queira que p, ou suspeite que p. Em vista desses comparativamente poucos exemplos, parece implausível sustentar que posso um conhecimento privilegiado, incorrigível de todos os diferentes estados mentais nos quais se poderia conceber que estivesse. Enquanto minha declaração sincera é suficiente para garantir a verdade do relato sobre a experiência que estou tendo no caso da perfuração dentária, algumas vezes pode ser nosso comportamento, e o contexto no qual ele ocorre, que serve como critério, como medida, segundo o qual a atribuição de um estado mental intencional para mim, até mesmo por mim próprio, que torne acessível a correção do relato.

### **As perspectivas de primeira e terceira pessoas e a busca de uma abordagem unitária da mente**

Estamos agora em melhor posição para distinguir as perspectivas de primeira e de terceira pessoas uma da outra. No

\*N. de T.: Em português: *A natureza de mentes verdadeiras*.

caso de conhecimento do que ocorre no e para o meu corpo, assim como corpos e objetos materiais em geral, meu conhecimento é, essencialmente, a partir de uma perspectiva externa, de terceira pessoa. Neste caso, esta é a única perspectiva que me é disponível, ou, de fato, a qualquer outra pessoa. Meu conhecimento dos seus estados mentais é também feito a partir de uma perspectiva de terceira pessoa, porque ele se baseia externamente no comportamento público que você exibe, assim como o seu conhecimento de meus estados mentais devem também ser a partir de um ponto de vista de terceira pessoa. No entanto, ao menos em alguns casos, parece que sei o que está se passando em minha mente diretamente, assim como você sabe o que se está passando na sua, a partir de uma perspectiva interna, privada, de primeira pessoa, que não requer a observação ou referência a qualquer comportamento externo.

É importante não confundir as perspectivas de primeira e de terceira pessoas com atribuições de estados mentais de primeira e terceira pessoas. Normalmente, será verdade que os estados mentais que atribuo a mim mesmo são feitos a partir de uma perspectiva de primeira pessoa, como quando eu tenho uma dor ou um comichão. Igualmente, quando eu atribuo estados mentais a você, estes serão necessariamente a partir de uma perspectiva de terceira pessoa, uma vez que eu não tenho outra forma de tomar conhecimento de seus estados mentais a não ser, geralmente, por meio de suas ações e comportamentos. No entanto, haverá aquelas ocasiões em que sou temporariamente levado a adotar uma perspectiva de terceira pessoa sobre meus próprios estados mentais, uma vez que terceiras pessoas conhecem minha mente melhor do que eu, e que eu baseio minha atribuição de meus estados mentais naquela base externa, em vez de

em uma perspectiva de primeira pessoa não confiável (neste caso).

O fato de que estados mentais possam ser atribuídos a fundamentalmente dois tipos diferentes de modos – a partir de uma perspectiva de primeira pessoa, baseando-se na percepção interior, e a partir de uma perspectiva de terceira pessoa, fundamentando-se na observação de comportamento – é considerado por alguns filósofos como fonte de problemas. Colin McGinn faz uma boa apresentação do caso: se o significado de um termo mental é governado pelo modo segundo o qual ele é aplicado, então isto parece fazer com que uma abordagem unitária do significado de tais termos seja impossível.<sup>13</sup> Quando eu atribuo uma dor a mim, eu o faço com base na percepção interior, a partir de uma perspectiva de primeira pessoa. No entanto, quando você atribui a mesma dor para mim, você o faz a partir de uma perspectiva de terceira pessoa, com base em meus gemidos e contorções. Uma vez que o significado do termo não pode ser extraído das condições sob as quais ele é aplicado, e uma vez que essas condições são muito diferentes quando feitas a partir de perspectivas de primeira e de terceira pessoas respectivamente, o significado fica sendo puxado em duas direções opostas.

Uma dessas direções é representada pelo cartesianismo, e leva ao ponto de vista de que os estados mentais são de natureza intrinsecamente interior, privada e não-física, acessíveis somente à pessoa que possui tais estados mentais. A outra direção é representada pelo behaviorismo analítico, que tenta reduzir os estados mentais a padrões de comportamento público externo possível e real. No entanto, como pode um dado estado mental possuir naturezas tão inteiramente disparates? Parece que somos forçados a escolher entre uma opção ou outra, uma vez que a formação de um conceito unitário do men-

tal parece estar descartada. No entanto, a completa redução a somente uma dessas perspectivas conduz a uma visão distorcida da natureza da mente.

Uma estratégia possível, sugere McGinn, seria tomar cada tipo de estado mental em seus termos próprios, e então tentar determinar qual a perspectiva que melhor explica o estado em pauta. Alguns estados mentais, tais como compreender uma teoria, terão sua aplicabilidade amplamente governada a partir de uma perspectiva behaviorista. Outras, como as sensações, darão melhores respostas a uma posição cartesiana. Uma extensão desta abordagem, em conformidade com as recomendações de McGinn, seria tratar uma perspectiva como primária em relação à outra, e então tentar encontrar um meio plausível de conectar o uso primário do conceito com seu aspecto secundário.

Assim, por exemplo, os wittgensteinianos têm argumentado que, embora seja verdade que não atribuímos dores a nós mesmos com base na observação do nosso comportamento, nossas autoatribuições só são possíveis porque há sinais comportamentais disponíveis para terceiras pessoas contra as quais nossas atribuições de primeira pessoa podem ser testadas quanto às suas correções ou incorreções. Neste caso, a perspectiva de terceira pessoa tem sido considerada de importância capital para que se chegue à compreensão do fenômeno mental em questão.

Tem sido dito, no entanto, que esta abordagem é muito menos plausível no caso das sensações, porque, como já salientamos, parece que estas são autoatribuídas com base em uma percepção privada interior, e não em conexão com a observação do comportamento, ou em referência a ele. Esta é a razão pela qual Galen Strawson, em seu livro *Mental Reality*, argumenta:

"A palavra 'dor' é apenas uma palavra para certo tipo de sensação desagradável, considerada independentemente de suas causas e efeitos comportamentais ou outras publicamente observáveis" (ênfase minha).<sup>14</sup> Semelhantemente, James Hopkins tem salientado que "objetos privados possuem naturezas a espera de serem reconhecidas"<sup>15</sup>, uma afirmação à qual Howard Robinson deu corpo convincentemente em seu livro, *Perception*.<sup>16</sup> Robinson deseja sustentar que dores e outras sensações deveriam ser concebidas como objetos privados. Ele escreve: "Um sujeito não tem e não necessita de critérios públicos para estabelecer como algo se apresenta para ele em uma experiência deste tipo",<sup>17</sup> e, dando uma clara guinada no dito de Wittgenstein de que "Qualquer coisa que irá parecer correta será correta, o que somente quer dizer que aqui não podemos falar sobre o que será correto",<sup>18</sup> Robinson diz que, no caso em pauta, "necessariamente, o que quer que eu julgue ser correto é correto". Consequentemente, de acordo com Robinson, dores e outras sensações podem ser reidentificadas por aqueles que as possuem na privacidade de suas próprias mentes, e assim pode-se falar a respeito delas com total independência dos critérios de comportamento público. Robinson apresenta um argumento convincente em favor desta posição. Ele nos solicita imaginar um caso em que alguém sofre uma dor na sua mão direita, e uma dor em sua mão esquerda ao mesmo tempo. Parece agora completamente crível que, puramente a partir do modo como essas dores são sentidas, a pessoa possa dizer se elas são, ou não, ambas instâncias do mesmo tipo de dor. Mas, se a identificação de uma dor como sendo semelhante a uma outra dor em um momento dado – identifica-

\*N. de T.: Em português: *Percepção*.

ção sincrônica – é plausível, então também devem ser situações nas quais duas dores sentidas em tempos diferentes são julgadas qualitativamente iguais. Neste caso, sustenta Robinson, a identificação ao longo do tempo de dores como sendo do mesmo tipo recorrente – identificação diacrônica ao longo do tempo – deve ser possível.

No entanto, no que diz respeito a qualquer tipo particular de estado mental, talvez deva ser dito que não pode haver nenhuma garantia prévia de que haverá um modo aceitável de minorar a tensão entre perspectivas de primeira e terceira pessoas, e isso, pensa McGinn, “pode de fato colocar um obstáculo permanente ao modo de se chegar a uma concepção de mente que seja teoricamente satisfatória”<sup>19</sup>. O mesmo problema, felizmente, não ocorre no caso de objetos materiais e de eventos, uma vez que se chega ao conhecimento deles a partir de uma perspectiva puramente de terceira pessoa. Devo, no entanto, por uma questão de espaço, deixar que o leitor prossiga seu estudo das questões suscitadas nesta seção, por sua própria conta.

### **O caráter holístico dos estados intencionais e sua relação com o estado de consciência**

Suponha que você tem uma dor nos ombros. O fato de que você está tendo a experiência dessa sensação não tem implicação alguma com o fato de você estar ou não tendo uma sensação de formigamento no seu pé. Suponha, no entanto, que você está tendo uma sensação de formigamento no seu pé bem como uma dor no seu ombro. A ocorrência dessas duas experiências, por sua vez, não tem implicação alguma quanto a você estar tendo

uma dor de dentes. Tanto você poderia estar ou não. Poderíamos descrever essa ocorrência dizendo que as sensações são atomísticas: a ocorrência de uma sensação individual não tem implicações quanto à existência, ou a não-existência, de qualquer outra sensação, assim, neste sentido, as sensações podem existir soltas e separadasumas das outras.

Mas, agora, pense em um estado de não-sensação, intencional, tal como a intenção de cancelar um cheque. Contrastando com as sensações, este estado mental certamente tem implicações para a existência de outros estados intencionais. Por exemplo, a intenção implicará todos os tipos de crenças de minha parte – a existência de um sistema monetário e de bancos, da instituição de fazer promessas, que há uma boa razão pela qual o cheque deva ser cancelado nesta ocasião, e assim por diante. Contrastando com o que é verdade para as sensações não-intencionais, não faria sentido supor que minha intenção de cancelar um cheque poderia existir isoladamente, sem uma vasta rede de outros estados intencionais. Um estado intencional pode ser comparado a uma peça de quebra-cabeças. Uma peça deve a sua identidade como uma peça às outras peças às quais ela está relacionada. Caso todas as demais peças do quebra-cabeças fossem destruídas, a peça individual perderia a sua identidade, pois não mais poderia desempenhar o papel de constituir parte de um quebra-cabeça. A peça teria se tornado simplesmente uma peça de cartão colorido com uma forma bizarra. Semelhantemente, a noção de um estado intencional isolado inteiramente sozinho é incoerente, porque qualquer estado individual, por sua própria natureza, deve pressupor a existência de muitos outros estados relacionados. Estados intencionais são como peças de quebra-cabeças quanto a outros aspectos importantes. É

da natureza das peças que elas possam encaixar-se para formar um todo coerente, mas é igualmente possível que um engano possa ser cometido na tentativa de colocar uma peça em uma posição na qual ela não se enquadre. Tentar encaixar a peça à força simplesmente entortaria ou distorceria o resto do quebra-cabeça. A capacidade dos estados intencionais de encaixar-se para formar um todo harmonioso, ou de serem inconsistentes entre si, bem como pressupor e implicar a existência de outros atos intencionais, é referido pelos filósofos como seu caráter holístico.

O estado de consciência desempenha um papel vitalmente importante na formação de um conjunto coerente de atitudes proposicionais. Com relação a essas atitudes, o estado de consciência significa estritamente autoconsciência, a habilidade de reconhecer que se tem determinadas crenças, juntamente com outras atitudes proposicionais, tais como desejos. O ponto crucial é que, a menos que as pessoas pudessem tomar ciência de suas próprias crenças, isto é, pudessem saber que as têm e pudessem refletir sobre elas, elas não teriam condições de localizar ligações lógicas, bem como inconsistências entre elas. No entanto, sem ser capaz de ver que adotar certas crenças logicamente leva à adoção de outras determinadas crenças, a inferência lógica não seria possível. Tampouco seria possível a alguém detectar inconsistência entre crenças distintas, e isso impediria aquela pessoa de ajustar racionalmente suas crenças, mantendo algumas e rejeitando outras, a fim de colocá-las em harmonia umas com as outras. Além disso, as crenças são racionalmente julgadas não somente pela comparação com outras crenças, mas procurando ver se a evidência disponível faz com que sua manutenção seja razoável. Ao que tudo isto aponta é que, a menos que uma pessoa pudesse

trazer à mente suas crenças e as examinar autoconscientemente, a capacidade de pensar racionalmente não seria possível.

Este é um resultado interessante e surpreendente, que auxilia na demonstração da vantagem evolucionária trazida pelo estado de autoconsciência. O estado de autoconsciência faz com que a racionalidade seja possível, e é somente por meio do exercício da racionalidade que os seres humanos têm tido a capacidade de compreender e explicar por que as coisas são como são. Juntamente com essa capacidade vieram os meios para manipular e controlar nosso entorno de um modo que supera amplamente tudo que animais não-humanos possam fazer, assim ampliando a possibilidade de nossa sobrevivência em face das mudanças hostis do meio ambiente. Por exemplo, caso um grande asteróide estivesse prestes a se chocar com a Terra e ameaçar a extinção da vida, animais não-humanos não poderiam fazer nada quanto a isso. Em contraste, não somente estamos cientes da ameaça e de por que ela existe, mas a nossa compreensão científica deste assunto está perto de nos fornecer a capacitação para diminuir ou evitar o perigo.

Mais importante ainda é a capacidade distintamente humana de construir linguagens e sistemas simbólicos que constituíram a formação de atitudes proposicionais e fizeram com que o exercício da racionalidade fosse possível. Isso, por sua vez, constitui a condição necessária à origem e à evolução dos arranjos sociais e políticos que distinguem o mundo humano daquele do animal, como Aristóteles já percebeu há muito tempo:

A razão pela qual o homem é um ser voltado à associação política em um grau mais elevado do que as abelhas ou outros animais gregários jamais poderão se associar é evidente. A natureza, de acordo com a nossa teoria, não faz nada em vão,

e somente ao homem foi dada a faculdade da linguagem. A mera emissão de sons serve para indicar prazer ou dor, e é, portanto, uma faculdade que pertence aos animais em geral: a natureza deles os permite atingir o ponto no qual têm a percepção de dor e prazer, e podem significar tais percepções uns para os outros. Mas a linguagem serve para declarar o que é vantajoso e o que é contrário, e, portanto, serve para declarar o que é justo e o que é injusto.<sup>20</sup>

### Há um traço do mental?

#### **EXERCÍCIO 1.8**

À luz da ampla discussão que compõe este capítulo, você pensa que é possível enumerar as condições que fazem a demarcação do mental e do não-mental? Há alguns elementos deste critério que são mais básicos do que outros? Tente dizer o que distingue o mental em relação ao não-mental e depois leia a minha argumentação a seguir.

Em primeiro lugar, coisas físicas, e talvez entidades abstratas, tais como números, podem existir independentemente de termos ou não ciência deles. Em contrapartida, sensações tais como dores e comichões existem se, e somente se, uma pessoa ou ser está atualmente consciente delas. Para ilustrar este ponto, um objeto físico, tal como uma árvore, pode existir independentemente de alguém a ver ou tocar, mas uma dor da qual ninguém tem ciência é certamente uma contradição nos termos. A importância do estado de consciência no que diz respeito ao mental não pode ser superestimada. Até estados intencionais que não envolvem sensação, tais como crenças e desejos, que podem existir temporariamente na ausência de um estado de consciência, devem em úl-

tima instância ser acessíveis ao estado de consciência para que possam ser genuinamente atribuídos a uma pessoa.

O estado de consciência é vitalmente importante também para assegurar coerência entre as atitudes proposicionais que abarcam estados intencionais que não envolvem sensação, porque a menos que se possa ficar ciente das atitudes proposicionais que se adota, não haverá condições para que se detecte e elimine qualquer inconsistência entre elas. A menos que este tipo de ajuste possa ser feito, a racionalidade é impossível. Resumindo, a conclusão à qual essas observações aportam é a de que o estado de consciência, e a autoconsciência, a sua manifestação mais elevada, são, em seus diferentes modos, absolutamente integrais à existência de uma mente humana em sua plenitude.

Em segundo lugar, é difícil entender como estados intencionais poderiam vir a ser de natureza puramente física. Como podem estados físicos, tais como acontecimentos neurofisiológicos no cérebro, serem dirigidos a um conteúdo que, em alguns casos, não tem, fora da mente, correspondência alguma no mundo? Além disso, os estados intencionais são essencialmente holísticos, mas parece não haver contrapartida deste traço em estados de coisas puramente físicas.

Em terceiro lugar, os fenômenos físicos só são acessíveis a partir de uma perspectiva de terceira pessoa, enquanto estados mentais podem ser acessados e atribuídos a partir tanto do ponto de vista de uma primeira pessoa – “interna” – e de uma terceira pessoa – “externa”. Levando isso mais adiante, o acesso e o conhecimento que as pessoas têm de seus próprios estados mentais pareceriam ser privilegiados, no sentido de que elas podem dizer, sem observação ou inferência, o que estão pensando ou sentindo, e, frequentemente, são as autoridades máximas quanto a

seus estados mentais. Em contrapartida, uma pessoa é capaz de atribuir estados mentais a uma outra somente com base em inferências feitas sobre observações do comportamento mostrado em público por aquela pessoa. Esses traços de estados mentais, sua privacidade radical, e a sua possibilidade de serem atribuídos a partir de perspectivas tanto de primeira quanto de terceira pessoas, fazem com que seja difícil ver-se como o mental possa alguma vez ser reduzido ao físico, ou eliminado em seu favor.

Finalmente, embora eu não tenha dito nada a esse respeito aqui, pode ter ocorrido a você que mais uma diferença importante entre o mental e o físico é que aquele pareceria ter um caráter não-espacial, enquanto é da própria essência dos fenômenos físicos que eles ocorram em regiões do espaço – veja, por exemplo, os campos magnéticos e gravitacionais –, e em muitos casos eles são os ocupantes literais daquelas regiões, como é o caso de objetos físicos, como árvores, planetas e partículas subatômicas. No entanto, qualquer discussão quanto à impossibilidade de se atribuir literalmente localização espacial a estados mentais, e somente com base nisto, portanto, não poderem ser identificados com qualquer estado físico de coisas, será posto até o Capítulo 3, seção “O mental e o espacial”.

### **Questões para reflexão**

1. Qual é o significado do termo “qualia”?
2. O que é intencionalidade? Qual é a diferença entre intencionalidade intrínseca e derivada ou intencionalidade “como se”? Essa distinção pode ser mantida?
3. Os estados mentais são privados?

4. O que você entende por perspectivas de primeira e de terceira pessoa?
5. “Descreva três características de estados mentais que, diz-se, os distinguem de estados físicos. Indique o ponto de vista de que as características de estados mentais fazem com que seja impossível que tais estados se encontrem dentro do mundo físico” (Associated Examining Board, questão nível A, 1997).
6. Por que se diz que o estado de autoconsciência seja necessário para que a racionalidade seja possível?
7. É possível produzir uma definição sem furos do mental que possa ser usada para distingui-lo do físico?

### **Sugestões para leituras aprofundadas**

- Uma excelente introdução e livro de referência é dado por GUTTENPLAN, S. ed., *A Companion to Philosophy of Mind*, Oxford, Blackwell, 1994. O capítulo de abertura “Um ensaio sobre a mente” é altamente recomendado por sua lucidez e interesse.
- GRAHAM, George. *Philosophy of Mind: An Introduction*. Oxford, Blackwell, 1993 também fornece uma estimulante e divertida introdução ao problema da mente/corpo e questões a ele relacionadas.
- HEIL, John. *Philosophy of Mind: A Contemporary Introduction*, London, Routledge, 1998 fornece um panorama claro desta área da filosofia.
- Um livro mais pesado, mas que vale muito a pena, é CARRUTHERS, Peter. *Introducing Persons*. London,

Croom Helm, 1986. Um elemento útil do livro é sua apresentação formal de uma variedade de argumentos que o leitor é estimulado a encontrar por conta própria.

- SEARLE, John. *Minds, Brains and Science* [Reith Lectures, 1984], Harmondsworth, Penguin, 1989 fornece uma introdução estimulante, acessível e provocante às principais questões envolvidas no problema mente/corpo.

- McGINN, Colin. *The Character of Mind*, Oxford, Oxford University Press, 1982 suscita e discute com profundidade e sofisticação muitas das importantes questões da filosofia da mente, mas as discussões são surpreendentemente condensadas, o que faz com que este livro não seja para quem está recém-iniciando.

## NOTAS

1. Extraído de uma brincadeira de Dickens, incorporado nas palavras da Sra. Gradgrind em *Hard Times*: "Eu acho que há uma dor em algum lugar na sala, mas não poderia positivamente assegurar que eu a tenho". Hume teria discordado: "Tudo aquilo que é distinto é distinguível; e tudo aquilo que é distinguível é separável por meio do pensamento ou da imaginação. Todas as percepções são distintas. Elas são, portanto, distinguíveis e separáveis, e não podem ser concebidas como existindo separadamente, e podem existir separadamente, sem qualquer contradição ou absurdidade". Apêndice A de: SELBY-BIGGE, L.A. ed. *Um Tratado da Natureza Humana*. Oxford, Clarendon Press, 1978, p.634. O que Hume quer dizer por uma percepção é uma experiência, isto é, um item mental. Portanto, ele está afirmando que experiências podem existir soltas e separadamente sem dependência de algo mais.
2. LEIBNIZ, G.W. *Monadology*. In: ERDMANN, J.E., ed. *Leibniz, Opera Philosophica*, 2.v., Berlin, 1940 [primeira edição póstuma].
3. Admite-se que tem havido alguns casos estranhos e aberrantes, como o relatado por Michael Lockwood em seu livro *Mind, Brain and the Quantum*, Oxford, Oxford University Press, 1989, em que um sujeito hipnotizado, a quem se solicitou manter sua mão esquerda em um balde de água gelada, negou, com aparente sinceridade, estar sentindo dor; no entanto, quando lhe foi dado papel e lápis, começou a escrever com sua mão direita que estava sentindo uma terrível dor. Este caso se parece com o de ocorrências de "visão cega" [blindsight], termo cunhado pelo psicólogo Larry Weisskrantz. Oliver Sacks descreve um caso de visão cega em seu fascinante livro, *An Anthropologist on Mars*, London, Picador, 1995. Um paciente, chamado Virgil, sofreu dano cerebral devido a uma falta de oxigênio e, posteriormente, insistiu que estava cego e não podia ver coisa alguma. No entanto, ele apanhava objetos e evitava obstáculos – em resumo, ele agia exatamente como se estivesse vendo, mas negava com sinceridade que tivesse qualquer consciência de estar vendo. É difícil saber como explicar tais casos; parece que sinais são recebidos pelos olhos e dão resposta, mas que nada é filtrado por meio da consciência da pessoa. Sugiro que podemos levar em consideração a existência destes casos altamente fora do comum precisamente porque não são a norma e não perturbam o uso normal dos conceitos de sentir uma dor, ou de ver coisas. Caso a situação fosse alterada, e a visão cega e dores não sentidas se tornassem predominantes, e as exceções passassem a ser o que agora consideramos como a visão e sensação de dor normais, então nossa noção do que seja sentir dor e ver sofreria uma revolução conceitual para acomodar esta mudança.
4. Deixo de lado as complicações suscitadas por conceitos tais como os de repressão, estados mentais inconscientes, autoengano e outros, por uma questão de espaço.
5. Alguns filósofos, de modo controverso, tentaram interpretar as causas das sensações como sendo o que as sensações representam, a respeito de que elas são. Veja, por exemplo, TYE, Michael. *Ten Problems of Consciousness*. Cambridge, MIT Press, 1995, c.4.

6. BRENTANO, F. *Psychology from an Empirical Standpoint*. [1874], [PANCURELLO, A.C. e TERRELL, D.B., trads.] New York, Humanities Press, 1973, p.88.
7. HEIL, J. *The Nature of True Minds*. Cambridge, Cambridge University Press, 1992, p.152.
8. HUME, D. *A Treatise of Human Nature*. L 1, iv parte, p.234.
9. Compare com: KENNY, A. *The Metaphysics of Mind*. Oxford, Clarendon Press, 1989, p. 39.
10. PROUST, M. *À la recherche du temps perdu*. [MONCRIEFF, C.K.Scott, trad.], London, Chatto and Windus, 1913-21, p.123-4.
11. CONRAD, J. *Lord Jim*. New York, Bantam Books, 1957, p.5.
12. HEIL, J. *The Nature of True Minds*. p. 154-5.
13. McGINN, C. *The Character of Mind*. Oxford, Oxford University Press, 1082, p.6-7.
14. STRAWSON, G. *Mental Reality*. Cambridge, MIT Press, 1994, p.245.
15. HOPKINS, J. Wittgenstein and Physicalism. *Proceedings of the Aristotelian Society*, 75, 1974-5, p. 121-46.
16. ROBINSON, H. *Perception*. London, Routledge, 1994.
17. Ibid., p. 104.
18. Ibid., p. 105.
19. McGINN, C. *The Character of Mind*, p. 7.
20. ARISTOTLE, *Politics*, 1253<sup>a</sup> 8-13, in: SMITH, J.A. e ROSS, W.D. eds. *The Works of Aristotle*, Oxford, Clarendon Press, 1910-52.

# 2

## DUALISMO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Ser capaz de descrever os traços principais do dualismo e como ele ocorre nas filosofias de Platão e de Descartes.
- Ser capaz de expor e criticar os argumentos em favor do dualismo construídos por Platão e por Descartes.
- Entender como Descartes tenta explicar a interação entre mente e corpo e ser capaz de avaliar a tentativa dele.
- Ser capaz de explicar as principais objeções ao dualismo, incluindo, sobre tudo, a falácia do homúnculo e o problema colocado pela suposição de que as almas podem ser distinguidas umas das outras e contadas.

### Dualismo substancial

A afirmação básica do dualista é a de que o ser humano é um composto de duas entidades distintas: uma alma ou mente, que é não-física, e um corpo físico. A pessoa – aquele que dispõe de consciência e de experiência em geral – é idêntico à alma, não ao corpo. Ao longo de sua vida, o corpo e a alma estão, de alguma forma, ligados um ao outro, de modo que um pode afetar o outro, mas, na morte, esta ligação é rompida. O corpo, então, sofre uma dissolução, mas a alma continua a existir indepen-

dentemente do corpo, por si só. Por fornecer elementos para a continuidade da existência de pessoas após a morte em uma forma não-física, desencarnada, o dualismo tem se mostrado uma doutrina atraente para muitas pessoas, especialmente para aquelas que seguem algum tipo de religião. A morte física não é o nosso fim: na condição de consciência pura, continuamos a existir em um mundo além deste aqui. Portanto, não há razão para que se tema a morte, porque ela não resulta em aniquilamento, mas simplesmente marca uma transição de um estado de ser para outro.

Concebida como uma entidade genuína, que pode existir por si só, independentemente de qualquer outra coisa, diz-se tradicionalmente que a alma é uma substância lógica, daí o termo "dualismo substancial". O termo "substância", quando encontrado pela primeira vez em um contexto filosófico, pode levar a enganos, e para que você entenda o que os filósofos querem dizer com isso você deve primeiramente livrar sua mente de usos tais como os ilustrados pela seguinte frase: "Que tipo de substância é esta que está em seu tubo de ensaio?". Nesse sentido, pensa-se em substância como sendo algum tipo de material ou coisa, tal como um plástico, cola ou água. O significado filosófico do termo não é de forma alguma parecido com isso, como um exemplo ajuda a esclarecer. Vejamos as seguintes linhas do poema "Morning at the Window"<sup>\*</sup>, de T. S. Eliot:

As ondas marrons de neblina lançam-se  
a mim  
Rostos esquisitos do fundo da rua,  
E lágrima de uma passante com saias  
enlameadas  
Um sorriso a esmo que paira no ar  
E desaparece ao longo do nível dos  
telhados.<sup>1</sup>

Agora, embora possamos facilmente conceber a pessoa que passa portando ou não um sorriso a esmo em seu rosto, não podemos conceber o sorriso separadamente da transeunte. Por que isso é assim? Claramente porque um sorriso não é nada mais do que uma disposição – o arranjo, se preferir – de partes do rosto da transeunte. Logicamente um sorriso não pode existir sem alguém que sorria, e a identidade do sorriso – o sorriso em particular, está entre muitos sorrisos possíveis – está ligada logicamente à identidade da

pessoa que está fazendo o ato de sorrir. A transeunte, por outro lado, tem o status de uma substância lógica, uma coisidade genuína; porque a sua existência não depende logicamente de alguma outra coisa, como acontece com o sorriso. A palavra "sorriso" é certamente um substantivo, e muitos de nós fomos ensinados na escola que os substantivos adquirem seus significados por meio da nomeação de pessoas, lugares ou coisas; assim poderíamos ser tentados a pensar, equivocadamente, que um sorriso fosse uma substância.

O antigo filósofo grego Aristóteles (384-322 a.C.) definiu uma substância como sendo "aquilo que não é dito de um sujeito, nem em um sujeito, por exemplo, um homem individualmente ou um cavalo".<sup>2</sup> Essa definição dá a ideia de que substâncias podem existir por si mesmas, contrastando com o que pode ser dito **de** elas, ou estar **em** elas. A distinção entre estar **em** e ser **dito de** pode ser explicada do seguinte modo: sorrisos qualificam-se como "particulares", isto é, como itens individuais. Por exemplo, pode haver um sorriso de João, um sorriso de Maria, e assim por diante, e são sorrisos **em** sujeitos. Os sorrisos devem sua identidade e existência às substâncias nas quais se encontram: eles são particulares logicamente dependentes. Diferentemente de um sorriso, a cor de uma face – digamos, vermelha – não se qualifica como um particular. Em vez disso, ela é um universal ou uma propriedade, isto é, um traço completamente geral, que pode ser verdadeiro de muitas coisas individuais ou de substâncias: o açougueiro local de cara vermelha, por exemplo, ou meu rubro tio Arthur. Propriedades contrastam com substâncias. Propriedades são possuídas por substâncias e não podem existir soltas e separadas por si próprias. Você não poderia encontrar uma

\*N. de. T.: Em português: "Manhã à janela".

instância de vermelho flutuando por aí independentemente, por si própria: melhor, alguém ou algo deve possuir a propriedade de ser vermelho.

A distinção entre substâncias e propriedades nos permite compreender como pode haver uma variante fraca do dualismo, qual seja, o dualismo de propriedade.<sup>3</sup> O dualismo de propriedade evita a noção de substâncias almas imateriais e sustenta, por outro lado, que o cérebro, uma coisa física composta, possui fundamentalmente diferentes tipos de propriedades – elementos mentais e consciência em particular – que são de natureza não-física, bem como propriedades físicas familiares, tais como massa, cor e textura. Terei muito mais a dizer sobre o dualismo de propriedade no Capítulo 6.

### Dualismo platônico

O filósofo grego Platão (c.427-347 a.C.) foi um dos principais expoentes do dualismo. Embora ele não tenha inventado a teoria, ele foi, tanto quanto sabemos, o primeiro a apresentar argumentos escritos a favor dela. Platão acreditava não somente na sobrevivência à morte por meio da alma, mas também na sua preexistência antes do nascimento. Naquele estado, supostamente ela estava familiarizada com as Formas – essências ou arquétipos perfeitos, atemporais, imutáveis, que, de acordo com Platão, eram os padrões originais ou moldes das coisas. Elas existem em seu próprio domínio acima do mundo dos sentidos, e só podem ser apreendidas pelo intelecto após um rigoroso treinamento em matemática e filosofia.

Platão afirmava que era por meio da sua participação nas Formas que as coi-

sas comuns, imperfeitas, mutantes e pecíveis, que se encontram à nossa volta, tornam-se os tipos de itens que são. Assim como as muitas coisas individuais que exibem beleza, por exemplo, há a suposição de haver a Forma do Belo, a própria Beleza, em virtude da qual as coisas imperfeitas do mundo dos sentidos são feitas belas. A alma, de acordo com Platão, se parece com as Formas, por ser “divina, imortal, inteligível, uniforme, indissolúvel, invariável e constante em relação a si própria”.<sup>4</sup> Essa descrição da alma foi extraída de *Phaedro*, obra na qual Platão apresenta uma série de argumentos em uma tentativa de demonstrar a existência da alma. Ao longo de muitos desses argumentos está razoavelmente claro que Platão concebe a alma como uma entidade não-física, e a pessoa é identificada com a alma. No entanto, Platão não é sempre consistente quanto às várias coisas que diz a respeito da alma. Por exemplo, em um lugar em seu diálogo *The Republic*, ele fala da alma como “o princípio da vida”, quase como aquilo que faz com que os seres vivos vivam. Neste sentido da palavra, qualquer coisa que esteja viva tem uma alma, por definição, pois “alma” é meramente sinônimo de “estar vivo”. Assim sendo, a alma não pode ser pensada como uma entidade, uma pessoa individual ou agente, contrariamente ao que Platão diz em outros lugares. Em *Phaedro*, Platão também diz que a única atividade intrínseca à alma, inseparável dela, é a razão pura. Somente esta sobreviverá quando a alma se separar do corpo por ocasião da morte. No entanto, não está claro que a sobrevivência do intelecto signifique a sobrevivência da pessoa.

Platão tinha muitos argumentos para sustentar o dualismo substancial, mas es-

\*N. de T.: Em português: *Fedro*. Lisboa: Guimarães Editores, 1994.

\*\*N. de T.: Em português: *A República*. Lisboa: Calouste Gulbenkian, 2007.

colherei somente um, extraído de seu diálogo, *The Alcibiades I*<sup>5</sup>, para análise. A razão pela qual escolho examinar este argumento é que ele incorpora um efeito particularmente sedutor, cuja exposição nos poupará de numerosos outros enganos que poderíamos cometer ao filosofar sobre a natureza da mente.

Antes de dar detalhes do argumento particular em questão, será necessário dizer, entretanto, um pouco mais sobre o que é um argumento, e como argumentos podem ser avaliados quanto à sua solidez ou falta desta. Neste livro, discutem-se dois tipos de argumentos: induktivos e dedutivos. Uma vez que os argumentos dedutivos serão discutidos detalhadamente no Capítulo 9, seções "Introdução" e "Indução e o argumento por analogia", aqui não direi nada mais a seu respeito; nossa atenção, por enquanto, será voltada somente ao argumento dedutivo.

Na vida cotidiana, a palavra "argumento" geralmente significa ter uma altercação, engajar-se em uma disputa, frequentemente acalorada. O sentido aqui dado à palavra não é esse. Essencialmente, em filosofia, um argumento consiste na tentativa de passar, racionalmente, ou logicamente, a partir de certas afirmações, que chamamos de "premissas", ao estabelecimento da verdade de uma outra, que chamamos de "conclusão". Em um argumento dedutivo isso é válido, a conclusão decorre logicamente das premissas. Isto é, se você afirma (coloca) a premissa, mas nega a conclusão, você estará se contradizendo formalmente. Contradizer a si próprio é uma perda de tempo a ser evitada. Seria como dar dinheiro a alguém com a mão direita e tomá-lo de volta com a mão esquerda – não se obtém coisa alguma. Para esclarecer o que comporta um argumento dedutivo válido, veja o seguinte exemplo simples:

Premissa 1: todas as mulheres são mamíferos.

Premissa 2: todos os mamíferos têm sangue quente.

Conclusão: portanto, todas as mulheres têm sangue quente.

Se você formula (1) e (2), você tem de aceitar que a conclusão se segue, sob pena de você se contradizer. É, no entanto, muito importante dar-se conta de que a validade não tem absolutamente nada a ver com a verdade ou falsidade das premissas ou da conclusão, mas depende inteiramente da estrutura do argumento. O argumento acima tem a seguinte estrutura:

Premissa 1: Todo A é B.

Premissa 2: Todo B é C.

Conclusão: Portanto, todo A é C.

Claramente, nesta representação da estrutura do argumento, não podemos dizer se as premissas e a conclusão são verdadeiras ou falsas, porque não dissemos o que são A, B e C. Isso ajuda a esclarecer que a validade depende somente da forma do argumento, e não a que ele realmente se refere. No primeiro exemplo acontece de as premissas e a conclusão serem verdadeiras. No entanto, veja:

Premissa 1: todos os porcos têm asas.

Premissa 2: nenhuma criatura com asas pode voar.

Conclusão: Portanto, nenhum porco pode voar.

O argumento é válido. A conclusão realmente não se deriva das premissas. Mas, enquanto a conclusão é verdadeira, acontece que ambas as premissas são falsas. O problema é que argumentos nos quais quer uma, quer todas as pre-

\*N. de T.: Em português: *Alcibiades*. Lisboa: Inquérito PEA, 1997.

missas sejam falsas não podem fornecer fundamento para que se pense que suas conclusões sejam verdadeiras, mesmo em casos nos quais aqueles argumentos sejam válidos.

### **EXERCÍCIO 2.1**

Há somente um modo confiável no qual é suficiente que você examine o argumento para assegurar-se de que a conclusão tem de ser verdadeira. Para você, quais condições devem ser apresentadas?

Há somente duas condições. Você pode dizer que a conclusão é verdadeira, e somente se:

1. O argumento é válido.
2. Todas as premissas são verdadeiras.

Os argumentos podem ter mais de duas premissas. Um argumento poderia ter centenas ou mesmo milhares delas. No entanto, ele só logra estabelecer sua conclusão se ele for válido e se todas as suas premissas, sem exceção, forem verdadeiras. O que eu disse até agora não parece levar a muita coisa, mas de fato coloca em nossas mãos uma ferramenta extremamente poderosa. Se alguém usa um argumento dedutivo em seu trabalho, e nós podemos extraí-lo e colocá-lo formalmente na forma de premissas e uma conclusão, estamos, então, em posição de verificar se ele logrou êxito, procurando ver se, em primeiro lugar, ele é válido, e em segundo, se todas as suas premissas são válidas. Caso o argumento falhe em um ou ambos os testes, então falhou no estabelecimento de sua conclusão. Isso não quer dizer necessariamente que a conclusão seja falsa: pode haver outros fundamentos independentes para pensar que ela seja verdadeira. Mas, se isso é o

caso ou não, ainda podemos dizer que o argumento não prova o que supostamente está provando – e isto é um ataque muito devastador ao argumento de qualquer um!

Todos os argumentos que delineei dizem respeito a classes de coisas, e este padrão de argumentos, que foi inventado por Aristóteles, é chamado de silogismo. Nem todos os argumentos dedutivos tomam esta forma, mas não dispomos de espaço suficiente para entrar em detalhes aqui (os leitores interessados podem investigar o tópico por conta própria), embora em capítulos subsequentes eu, de fato, solicito a vocês que avaliem alguns exemplos de argumentos dedutivos não-silogísticos.

Frequentemente pode ser bastante difícil extrair argumentos de textos filosóficos e colocá-los formalmente. Na próxima seção, começo expondo informalmente o argumento dado por Platão em seu diálogo *Alcibiades I* para a existência da alma, e depois forneço uma assertiva formal dele, à qual os convido a avaliar sozinhos, usando as diretrizes que dei anteriormente. (Ao avaliar o argumento, reportem-se aos detalhes de como proceder: minha própria discussão vem após a declaração formal do argumento.) Como você verá, este tipo de exercício ocorre ao longo do livro, e deveria levá-lo a pensar por si mesmo.

### **Um argumento platônico a favor do dualismo**

O argumento é assim: Sócrates pergunta a Alcebiades se ele usa a fala ao falar, com o que este prontamente concorda. Sócrates então apresenta uma quantidade de exemplos com o propósito de estabelecer a conclusão de que o usuário de uma coisa, por exemplo, um sapateiro, e a coisa que ele usa, digamos, uma

ferramenta de corte, são numericamente diferentes uma da outra. Novamente, Alcebíades concorda de bom grado. Sócrates salienta então que um homem não usa somente suas mãos, mas o corpo inteiro, e que, portanto, assim como uma pessoa é distinta da ferramenta que usa, ou de um instrumento que toca, assim segue-se que ele deve ser distinto de seu corpo. O argumento pode ser colocado mais formalmente e se desenvolve resumidamente do seguinte modo:

1. O usuário de uma coisa e a coisa usada são numericamente duas coisas diferentes e distintas, isto é, duas substâncias lógicas.
2. Uma pessoa usa seu corpo.
3. Portanto, uma pessoa deve ser numericamente diferente e distinta de seu corpo, ou deve ser uma substância lógica não-corpórea, isto é, uma alma.

Antes de prosseguir sua leitura, veja o Exercício 2.2.

### **EXERCÍCIO 2.2**

Quais críticas você pode fazer sobre este argumento? Trabalhando sozinho, escreva-as e depois discuta a questão com alguém mais.

A primeira premissa pode certamente e obviamente ser aceita como verdadeira, caso o utensílio utilizado seja tal como uma faca, ou um instrumento musical, tal como uma harpa ou um piano. No entanto, não obstante o fato de estarmos falando de pessoas utilizando seus membros e corpos, eles usam tais coisas no mesmo sentido em que usam um utensílio ou um instrumento? Podemos ter uma boa suspeita de uma jogada quanto à palavra "usar" aqui. Afinal, pode-se dizer a alguém

para não ser tão preguiçoso, mas usar sua mente – isto é, aplicar sua inteligência a um problema. Porém, deveria esta utilização ser tomada para estabelecer que uma pessoa é inteiramente diferente de sua mente? Até mesmo um dualista ficará sem resposta para essa questão, caso uma pessoa não possa ser identificada com sua mente ou alma.

Sugiro que necessitamos perguntar se as pessoas mantêm o mesmo tipo de relação para com seu corpo e membros, assim como para uma ferramenta que usam, ou um instrumento que tocam. Uma razão para pensar que esse não pode ser o caso é a seguinte: Quando eu uso um formão, eu faço isso o apanhando e segurando em minha mão. Mas quando eu uso a minha mão, eu não tenho de apanhá-la e segurá-la – eu a movimento simplesmente. Em outras palavras, eu posso mover a minha mão sem ter de fazer qualquer coisa antes. Eu não posso, por outro lado, simplesmente mover uma ferramenta diretamente. Tenho que fazer isso fazendo alguma outra coisa – agarrando-a em minha mão ou boca, soprando nela, golpeando-a sobre a mesa até que as vibrações façam com que a ferramenta passe a funcionar, e assim por diante. Eu não posso mover uma ferramenta do mesmo modo que move meus próprios membros. Isso sugere fortemente que eu não tenho com meu próprio corpo e membros a mesma relação que tenho com uma ferramenta que eu uso.

Prosseguindo nessa linha de pensamento, se ao fazer uso de um formão eu dou um escorregão e corro a minha mão, comumente descreveríamos o que aconteceu dizendo que eu me machuquei. Isso porque minha mão é uma parte de mim, em contraste com o formão, que simplesmente não é. Um dualista dedicado, por outro lado, teria de insistir que este modo corriqueiro de descrever o que aconteceu está errado. A descrição preferida do que

aconteceu, para o dualista, tem que ser que o dano foi causado ao meu corpo, mas não a mim, porque eu sou realmente uma alma incorporada. Além disso, coisas não-físicas, como a alma, não podem ser machucadas por objetos físicos, como um formão. No entanto, a insistência do dualista nessa caracterização do que aconteceu pode ser vista meramente como uma fuga da questão em favor do dualismo desde o princípio. A afirmação de que uma pessoa usa seus membros e corpo da mesma forma como usa ferramentas ou instrumentos é igualmente tendenciosa. Podemos reconhecer o que está errado no argumento de Platão uma vez que nos demos conta de que ele contém, de fato, um trocadilho ou uso equivocado da palavra "usar". Em seu sentido literal, o usuário de uma coisa e a coisa usada são duas coisas diferentes. No entanto, a menos que a verdade do dualismo seja assumida desde o início, uma pessoa e seu corpo não são duas entidades diferentes. Consequentemente, eu não uso minhas mãos ou meu corpo no mesmo sentido em que eu uso um instrumento. O movimento do instrumento é o resultado de algo mais que eu fiz antes, mas as ações das quais dependem os movimentos de minha mão e corpo, quando eu os move, não são o resultado de algumas ações anteriores que eu realizei. Certamente, não causa dano descrever alguém como usando seus membros e corpo, desde que não se perca de vista que isto representa um uso metafórico, e não literal, da palavra "usar".

O que está subjacente no argumento de Platão em *Alcibiades I* é o fato de que há inúmeros modos de se falar, nos quais um contraste de algum tipo pode ser estabelecido entre uma pessoa e seu corpo, e essas locuções fazem com que pareça que os dois devam ser coisas distintas e diferentes. Por exemplo, dizemos que uma

pessoa **tem** um corpo forte, não que ela **é** um corpo forte, e a partir disto poderíamos ser tentados a argumentar que, por haver casos nos quais o possuidor de uma coisa e a coisa possuída são genuinamente duas coisas separadas, como ilustra o uso "Patrício tem um potente carro", consequentemente o possuidor de um forte corpo deve ser numericamente distinto do corpo que é possuído. No entanto, deveria ficar claro que isto simplesmente reprisa o erro que Platão cometeu em seu dialogo com Alcebíades. Ao final de contas, em vez de dizer que alguém tem um corpo forte, podemos facilmente descrevê-lo, por outro lado, como sendo fisicamente forte, e esta forma de expressão não traz nenhuma sugestão de que a pessoa seja uma entidade distinta de seu corpo. Da mesma forma, tomando um exemplo emprestado de Anthony Flew, para descrever alguém como açoitando seu corpo em protesto na linha de chegada para ganhar a corrida, não quer dizer que a pessoa em causa deva ser uma alma, literalmente brandindo um relho incorpóreo, mediante o qual ela incita um corpo enfraquecido a fazer maiores esforços. Tal descrição é meramente um modo pictórico de dizer que o vencedor da corrida fez um grande esforço de vontade.

Como Gilbert Ryle (1900-1976) salientou em *The Concept of Mind*<sup>6</sup>, não devemos ser levados a tirar, a partir da "absurdidade de tais colocações", a conclusão de que uma pessoa deva ser descrita como uma associação entre um corpo e um não-corpo. Os exemplos mostrados nos permitiriam chegar a essa conclusão somente se estivéssemos preparados para aceitar que o seguinte argumento estabeleceu sua conclusão:

1. Expressões que fazem referência a pessoas e expressões que fazem referência ao corpo de pessoas não

\*N. de T.: Em espanhol: *El Concepto de lo Mental*. Barcelona: Paidos, 2005.

- são equivalentes quanto ao significado.
2. Quando uma expressão tem um significado diferente de outra expressão, isto só pode ser porque tais expressões referem ou nomeiam substâncias lógicas diferentes e distintas.
  3. Portanto, as pessoas devem ser substâncias lógicas, que são diferentes e distintas de seus corpos.

Agora tente o Exercício 2.3, antes de ler minha discussão sobre este argumento.

### EXERCÍCIO 2.3

As duas premissas no argumento são verdadeiras, ou pode uma delas, ou ambas, ser rejeitada? Lembre-se: uma vez que o argumento seja válido (como este o é), a conclusão não pode ser estabelecida se ao menos uma das premissas é falsa. Decida se as premissas são verdadeiras ou falsas, e escreva as suas razões. Depois discuta a questão com outra pessoa, antes de prosseguir a sua leitura.

Uma vez que o argumento seja válido, caso desejemos rejeitar a conclusão, devemos mostrar que uma ou ambas as premissas são falsas. Creio que podemos aceitar que a primeira premissa seja verdadeira, como os vários exemplos já dados demonstram: há vários casos nos quais falar a respeito de uma pessoa e falar a respeito de seu corpo não são equivalentes quanto ao significado. Assim sendo, caso haja alguma falha no argumento, esta deve estar na segunda premissa. Esta pode ser facilmente demonstrada que é falsa, e uma forma como isso pode ser feito é a seguinte. Faça a seguinte consideração: a expressão “estrela matutina” tem um significado diferente de “estrela vespertina”; no entanto,

essas duas expressões se referem a uma e à mesma entidade, qual seja, o planeta Vênus, e não a duas substâncias lógicas diferentes. Um tipo diferente de exemplo que também pode ser usado para expor o erro que se está examinando é fornecido pela afirmação “Eu vi Patrício à noite passada”, que claramente tem um significado diferente de “Eu vi o corpo de Patrício à noite passada”. O último uso implica que eu vi Patrício morto, o que não faz o primeiro. Mas, novamente, isso não pode possivelmente implicar que por haver visto o cadáver de Patrício eu deva, portanto, ter visto uma entidade que é diferente e distinta do próprio Patrício. Quando eu vejo Patrício vivo, vejo um organismo vivo, um ser humano em funcionamento. Ao ver Patrício morto, vejo meramente o mesmo organismo que cessou de funcionar. Não há nada nesses modos comuns de falar que possa estabelecer um dualismo, ou mesmo fazê-lo plausível. Isso é certamente verdade que comumente falamos de pessoas mortas como se houvessem partido, como se tivessem desertado seus corpos, que agora são conchas vazias. Mas estas formas de falar são simplesmente figuras de palavras, e só parecerão dar suporte ao dualismo, caso se esteja, antes, já convencido da verdade da doutrina.

### Dualismo cartesiano

Depois de Platão, o mais famoso expoente do dualismo foi René Descartes (1596-1650). O novo ímpeto que Descartes deu ao dualismo foi tão surpreendentemente original e atraente que a versão por ele formulada é chamada de dualismo cartesiano, em sua homenagem.<sup>7</sup> A agenda para o moderno debate sobre a natureza da mente foi, assim, estabelecida por Descartes em escritos tais como *The Discourse*

*on the Method*\* (publicado em 1637) e *Meditations*\*\* (publicado em 1642).

O dualismo cartesiano, à semelhança da variante platônica, afirma que uma pessoa é uma e a mesma coisa que uma alma incorporada, uma substância lógica imaterial destituída de todos os elementos de corpos materiais, e, particularmente, de extensão, pois a alma não tem nem comprimento, largura ou profundidade, e, consequentemente, não ocupa volume no espaço. Ela é, portanto, em princípio, indivisível, pois é impossível ter-se uma metade ou um sexto de uma alma. A invisibilidade da alma tem sido usada como uma prova alegada da sua imortalidade, com base no argumento de que somente coisas que podem se deteriorar em partes são perecíveis.

Concebida como uma alma, uma pessoa é totalmente distinta e diferente em espécie de seu corpo físico, que ocupa espaço e tem extensão. Descartes se refere à alma como uma *res cogitans*, uma coisa pensante. Nas *Meditations*, ele explicita melhor o que ele quer dizer com isto:

Pensar é um outro atributo da alma: e aqui descubro o que propriamente pertence a mim mesmo. Somente isto é inseparável de mim. Eu sou – eu existo: isto é certo, mas quão frequentemente? Tão frequentemente quanto eu penso; pois poderia até acontecer de, parando eu inteiramente de pensar, ao mesmo tempo pararia juntamente de ser. Eu agora não admito nada que não seja necessariamente verdade: eu sou, portanto, falando com precisão, somente uma coisa pensante.<sup>8</sup>

Esta passagem deixa claro que pensar não é meramente um traço que a alma tem por acaso, mas sua característica

essencial, ou essência, um traço que lhe é indispensável e faz com que ela seja o tipo de coisa que ela é. Assim, a alma não pode perder a sua essência sem, consequentemente, cessar de existir, eis porque Descartes estava preparado para abraçar a extravagante afirmação de que a alma humana está sempre pensando, mesmo no túmulo.<sup>9</sup>

Quando Descartes fala de pensar, está usando o termo de modo deliberadamente estendido, para incluir não somente atividades intelectuais, como a cogitação e o raciocínio, mas até mesmo a experiência de sensações de vários tipos, bem como estados emocionais, como raiva e ciúme. Descartes também menciona dúvida, compreensão, afirmação, negação, ter desejos ou não, imaginação e todas as formas de percepção sensorial como exemplos de pensar. Além disso, e crucialmente, para Descartes, o traço distintivo do mental é a consciência, e isso quer dizer que ele considerava a própria ideia de estados mentais inconscientes e não-conscientes como uma contradição em termos. A ênfase que Descartes coloca na consciência como a característica essencial da mente pode ser encontrada no seguinte comentário extraído de seu *The Principles of Philosophy*<sup>10</sup>: “Pelo substantivo ‘pensamento’ significo tudo que ocorre em nós, de modo que temos consciência dele, na medida em que é um objeto da consciência”.<sup>10</sup>

Em contraste com a alma, o corpo é *res extensa*, uma coisa cuja essência, em comum com outras coisas materiais, é extensão no espaço. O corpo é uma máquina complexa, não-consciente, capaz, diferentemente da alma, de divisibilidade em partes. Arranjos de matéria, embora complexos, são incapazes de pensamen-

\* N. de T.: Em português: *Discurso do método*. São Paulo: Martins Fontes, 2007.

\*\* N. de T.: Em português: *Meditações metafísicas*. São Paulo: Martins Fontes, 2005.

\*\*\* N. de T.: em português: *Princípios de filosofia*. São Paulo: Editora Hemus, 2007.

to e de percepção. Somente as mentes, concebidas como almas incorpóreas não-físicas, podem possuir elementos mentais. Os processos que ocorrem no interior do corpo podem ser explicados em termos científicos e mecânicos, e ao ver o corpo desse modo, Descartes estava lançando os fundamentos para uma abordagem científica do que faz com que coisas vivas vivam. Por outro lado, Platão, como vimos em *Republic*, algumas vezes afirmou que a alma é o princípio da vida, que é a alma que faz com que coisas vivas vivam (ver seção “Dualismo platônico”), e isto foi depois transmutado por alguns pensadores na crença de que os organismos vivos são animados por um misterioso *élan vital* espiritual, ou força-de-vida, que fica além da abrangência da investigação empírica. A virtude do modo como Descartes vê a matéria é que ele afasta o mistério e faz com que o funcionamento das coisas vivas possam, em princípio, ser explicadas em termos de química e física. O corolário da abordagem de Descartes é que o corpo não morre porque a alma se vai, como pensou Platão, mas, sim, que a alma vai embora porque o corpo parou de funcionar.<sup>11</sup>

Resumindo, para Descartes, a alma é consciente, mas incapaz de extensão, enquanto o corpo, e a matéria em geral, têm extensão, mas são intrinsecamente incapazes de ter consciência e mentalidade.

### **Animais e máquinas**

Somente os seres humanos têm alma, sustentava Descartes, e, consequentemente, os animais não têm pensamento ou consciência. Os animais, ou os brutos, como algumas vezes Descartes se referia a eles, deveriam ser vistos como autômatos, ainda que máquinas “incomparavelmente mais bem arranjadas e adequadas a movimentos mais admiráveis do que qualquer máquina inventada pelo homem”.<sup>12</sup> A úl-

tima parte deste comentário torna claro que Descartes pensa que é impossível estarmos causando sofrimento aos animais, quando os matamos e os comemos, porque estaríamos fazendo essas coisas tão somente a máquinas não-conscientes, e, portanto, não seríamos mais culpáveis de causar maus tratos do que quando mandamos um automóvel para o ferro velho.

### **Os dois testes de Descartes para a presença da alma**

Descartes acreditava que a ausência da alma nos animais poderia ser estabelecida por dois testes, que descreveu em seu *Discourse on the Method*. Em primeiro lugar:

[Os animais nunca poderiam] usar palavras ou outros signos construídos, como nós fazemos para declarar nossos pensamentos aos outros... [Em segundo lugar,] enquanto eles poderiam fazer muitas coisas tão bem, ou melhor, do que nós, eles falhariam, impreterivelmente, em outras, revelando que não agem segundo o conhecimento, mas somente a partir da disposição de seus órgãos. Pois enquanto a razão é um instrumento universal que pode servir em todos os tipos de circunstâncias, esses órgãos necessitam de um arranjo especial para cada ação particular; assim que é imoralmente impossível que uma máquina pudesse conter tantos arranjos variados para que pudesse agir em todos os eventos da vida do modo que a razão nos permite agir.<sup>13</sup>

O primeiro teste se atém ao fato de que nós podemos usar a linguagem, e os animais não. Descartes, antecipando a Chomsky no século XX, reconhece que a verdadeira posse da linguagem envolve a habilidade de tanto compreender quanto produzir uma variedade infinita de frases que não foram encontradas previamente,

e ser capaz de responder linguisticamente de modo oposto ao longo de toda uma gama de situações novas. Consequentemente, embora você, sem dúvida, nunca tivesse se deparado com as frases "Perfume de abraços o assaltaram todo ele. Obscuramente, com carne faminta, ele silenciosamente suplicava adorar" até as encontrar em *Ulysses*, de James Joyce, você compreendeu mais ou menos instantaneamente o que elas significavam. As respostas linguísticas do homem não se limitam a umas poucas rotinas estabelecidas, que poderiam ser reproduzidas invariavelmente em circunstâncias inteiramente inapropriadas, assim como papagaios que foram ensinados a dizer obscenidades, sem compreender o significado das palavras que reproduzem. Esta habilidade única e distintiva do ser humano, acreditava Descartes, era uma das operações "que dependem da consciência" e que "são somente próprias a nós como homens".<sup>14</sup> Uma vez que Descartes está preparado para atribuir consciência somente à alma, não é surpreendente que ele pense que somente a presença da alma pode explicar a habilidade linguística do homem.

O segundo teste de Descartes assemelha-se muito ao primeiro, mas parece ter sua aplicação voltada, por outro lado, a comportamento não-lingüístico. Os seres humanos demonstram uma adaptabilidade e criatividade em suas respostas a novas situações, muito maior do que no caso de animais.

Descartes então nos desafia a explicar como é que os seres humanos podem passar estes dois testes, enquanto animais e máquinas não podem, a menos que o dualismo seja verdade. Descartes pareceria estar afirmando que nunca seríamos capazes de fazer uma máquina que pudesse genuinamente construir e compreender frases. Seres puramente materiais,

independentemente de seu grau de sofisticação e complexidade, nunca poderiam, em princípio, fazer uso da linguagem e engajar-se em uma conversação genuína e discursar em geral. Fantasias de computadores e robôs que falam são simplesmente fantasias, se Descartes está correto, e do mesmo modo, caso algum cientista cognitivo esteja tentando construir uma máquina que possa genuinamente engajar-se em discurso, então ele está perdendo o seu tempo e esforço, segundo Descartes.

Antes de prosseguir sua leitura, veja o Exercício 2.4.

#### EXERCÍCIO 2.4

Quão convincentes são para você os dois testes de Descartes para estabelecer o dualismo? Para você, há alguma razão para supor que a habilidade linguística humana e o escopo de produção e respostas comportamentais só podem ser explicados por meio da postulação da presença de uma alma? Escreva as suas razões e, após discuti-las com mais alguém, prossiga a sua leitura.

A objeção essencial aos testes de Descartes foi feita por Baruch Espinoza (1632-1677), que escreveu: "Ninguém até agora estabeleceu limites aos poderes do corpo, isto é, ninguém ainda foi ensinado pela experiência o que o corpo pode realizar exclusivamente pelas leis da natureza, na medida em que ela é considerada como extensão".<sup>15</sup> O que diz Espinoza é inteiramente razoável. A sua implicação é a de que talvez um dia possamos vislumbrar uma explicação de como a aquisição e o uso da linguagem se tornam possíveis, que dispensará a necessidade de invocar misteriosos processos não-físicos. Em congruência com essa esperança, Noam Chomsky postulou que cada um de nós contém um dispositivo de



aprendizado de linguagem\* (DAL), e este, presumivelmente, consistiria de estruturas físicas altamente complexas e precisas, e de circuitos no cérebro. O desafio de descrever tal sistema, sem falar em seu modo de operação, é ainda um daqueles com os quais nos deparamos hoje, no entanto. Mas, em vista dessa dificuldade, desistir e concluir que o veículo de nossa capacidade de usar a linguagem deve de alguma forma ser encarnado nas fantasmagóricas vísceras paramecânicas da alma não ajuda em nada. Almas, caso existissem, devido à sua própria natureza, estariam além da abrangência da investigação empírica, e afirmar aquele processo no recôndito da alma somente pode ser invocado para explicar que a aquisição da linguagem parece simplesmente marcar um retrocesso ao mistério.

### A relação mente/corpo

As palavras no monitor à minha frente, à medida que digito esta passagem, estão aparecendo porque eu assim desejo, ou, dito de outra forma, meu desejo de produzir estas palavras está ajudando a fazer com que apareçam. Isto certamente parece ser uma explicação causal: um estado de mente, meu desejo, age como causa para fazer com que uma cadeia de eventos venha a existir, que inclui, entre outras coisas, o movimento físico dos meus dedos sobre o teclado, como um estágio intermediário a caminho da produção de palavras na tela. Neste caso, a causalidade se passa da mente para o mundo. Mas, do mesmo modo, as palavras na tela estão causando que eu as veja. Luz está vindo das palavras, entrando em meus olhos e estimulando a minha retina e nervo ótico, conduzindo a um estado mental, que consiste em minha percepção visual

das palavras. A cadeia de causas e efeitos, neste caso, vai do mundo para a mente. A menos que estejamos preparados para negar que a mente possa afetar o mundo físico, e, por sua vez, ser influenciada por ele, alguma forma de explicar como isso é possível deve ser encontrada. Uma vez que para Descartes os estados mentais pertencem à alma não-corpórea, ele deve encontrar algum modo de explicar como a alma pode fazer surgir eventos corpóreos, e como acontecimentos que ocorrem no corpo e para ele podem ser registrados na alma.

A solução de Descartes foi a de fazer a conjectura de que a alma opera por meio da glândula pineal, que se encontra no sistema límbico, no meio do cérebro. Supostamente, a alma dá à glândula pineal um pequeno estímulo, que é ampliado por uma cadeia de causas e efeitos físicos, dando origem a um movimento corporal, uma ação, tal como a minha digitação destas palavras. Descartes fazia a conjectura de que os nervos são como finos tubos contendo o que ele chamava de "espíritos animais". Estes não são fantasmas não-físicos de animais falecidos, mas completamente de natureza física, consistindo, supostamente, de sangue altamente rarefeito.

Por que Descartes escolheu a glândula pineal como a sede da alma? Havia duas razões principais. Em primeiro lugar, Descartes observou que a glândula pineal é muito leve e móvel. Isto fez dela um instrumento adequadamente sensível, que poderia responder a estímulos precisos da alma.

Em segundo lugar, Descartes acreditava que na percepção comum não percebemos objetos físicos externos, mas imagens na cabeça, que aqueles fazem surgir. As duas imagens vindas dos olhos supos-

\* N. de T.: Embora encontre-se a expressão "aquisição de linguagem", ela não é adequada, pois a linguagem não é obtida, conseguida ou, menos ainda comprada, ela é, efetivamente, aprendida.

tamente se fundiriam, e seriam projetadas na superfície da glândula pineal, como se tivéssemos uma TV em miniatura ou uma tela de cinema em nossos cérebros. A visão consiste na alma prestando atenção na imagem projetada na tela propiciada pela glândula pineal.

Antes de prosseguir, veja o Exercício 2.5.

### **EXERCÍCIO 2.5**

Descartes demonstra uma considerável ingenuidade ao descrever um mecanismo mediante o qual uma alma não-física, não-extensa pode afetar o corpo físico extenso, e, por sua vez, ser afetada por ele. Você pode ver algum problema nesta abordagem? Pense sonzinho sobre isto e depois discuta com uma outra pessoa.

A dificuldade que pode ter ocorrido a você é esta: como pode a alma, desprovida de todas as dimensões físicas, possivelmente afetar e ser afetada pelo corpo extenso? Esta parece ser uma dificuldade em princípio, que nenhuma dose de ingenuidade neurofisiológica pode contornar.

Na sua Sexta Meditação, Descartes faz outra tentativa de explicar a causação psicofísica. Esta tentativa teve suas raízes no descontentamento de Descartes com a glândula pineal como o único ponto de contato por meio do qual, supostamente, age a alma. Mesmo se concedemos, para o bem do argumento, que a interação pudesse ocorrer pela glândula pineal, o modelo que resultaria de como uma pessoa se relaciona com seu corpo estaria em desacordo com nossa experiência, como se deu conta Descartes. Quando sentimos dor, por exemplo, ou a sensação de fome, temos a sensação em questão no membro ferido ou na boca do estômago. Não a percebemos como se estivesse à distância meramente,

do ponto de vista intelectual e espacial. Uma analogia torna isto mais claro. Meu carro pode ser batido e danificado enquanto estou dirigindo, mas eu, pessoalmente, não sinto nada no lugar onde o dano ocorreu, embora da minha posição dentro do carro eu possa ver que a lateral, por exemplo, foi danificada. Do mesmo modo, se falta água no radiador, não é o mesmo que a sede que sinto quanto estou desidratado. Eu sei que estou desidratado porque experiencio a sensação de ter sede, mas eu só tomo conhecimento que a água do radiador do carro está baixa quando acende uma luz vermelha no painel, ou quando eu vejo vapor saindo do capô do carro. Se a alma interagisse somente por meio da glândula pineal, então parece que minha percepção de dano ao meu corpo, ou o fato de que estou com sede, teria de ser semelhante ao que acontece quando vejo que meu carro foi danificado por uma batida ou vejo a luz vermelha avisando que o nível de água do radiador está baixo. Eis por que Descartes, usando uma analogia semelhante à do carro, diz:

Eu não estou presente em meu corpo meramente como um piloto está em seu navio; eu estou mais firmemente ligado a ele, como se estivesse misturado a ele, de modo que ele e eu formamos uma unidade. Senão, quando o corpo é ferido, eu, que sou simplesmente um ser consciente, não sentiria dor devido a isso, mas perceberia o ferimento por um simples ato de compreensão, como o piloto percebe pela visão qualquer dano causado por quebra que possa haver no navio; e quando o corpo necessita de comida ou bebida, eu devo explicitamente compreender o fato, e não ter sensações confusas de fome e sede.<sup>16</sup>

Para superar este problema, Descartes propôs que sensações corporais, tais como dor, fome e sede, às quais chama de modos confusos de consciência, deveriam

ser explicados como oriundas do fato de a "mente estar unida ao corpo, como se estivesse misturada a ele".<sup>17</sup>

Esta alegada entremescla da alma com o corpo cria o que Descartes chama de uma "união substancial", e é essencial à nossa natureza humana:

Os seres humanos são feitos de corpo e alma, não pela mera presença ou proximidade de um com o outro, mas por meio de uma união substancial... Se um ser humano é considerado, é considerado em si mesmo, como um todo, ele é uma união essencial, porque a união que liga um corpo humano e uma alma um ao outro não é acidental a um ser humano, mas essencial, uma vez que um ser humano sem ela não é um ser humano.<sup>18</sup>

O que isto quer dizer é que não há a divisão precisa entre a mente e o corpo, que estavam implicadas nas afirmações anteriores de Descartes. As sensações parecem formar uma classe híbrida intermediária entre os puros atos de intelecção da alma e as ocorrências físicas no corpo. Tanto os traços físicos quanto os não-físicos das sensações supostamente estão de alguma forma integrados sem costura, e dessa forma compreendem a união substancial de que fala Descartes.

Agora veja o Exercício 2.6, antes de seguir adiante.

### **EXERCÍCIO 2.6**

Você pode ver algum problema na sugestão de Descartes de que a união substancial entre a alma e o corpo é constituída pela alma sendo "como se estivesse misturada com o corpo?". Trabalhe sozinho ou com alguém mais e escreva a sua resposta antes de continuar.

Deveria estar claro que há um sério obstáculo ao que Descartes está afirman-

do. Como pode a alma, carente de todas as dimensões, ser entremesclada com um corpo extenso? Contraste a entremescla da manteiga em uma mistura de bolo. O que acontece é que as partículas de manteiga se espalham na mistura e se colam às partículas que a compõem. Mas tanto as partículas de manteiga quanto a mistura ocupam um lugar no espaço; portanto, não há dificuldade alguma em determinar como elas podem ser entremescladas. Em última instância, então, a tentativa de Descartes de explicar como a mente e o corpo interagem e juntos formam uma união deve ser reconhecida como um fracasso.

### ***Outras tentativas de explicar a relação mente/corpo***

Duas bem conhecidas tentativas de circunscrever o problema da interação entre mente/corpo foram sugeridas por Nicholas Malebranche (1638-1715), a quem devemos a doutrina do ocasionalismo, e por Leibniz (1646-1716), que propôs a teoria do paralelismo psicofísico.

De acordo com o proponente do paralelismo, a alma e o corpo são como dois relógios aos quais Deus deu corda e sincronizou, a fim de que haja uma correlação apropriada entre eventos mentais e corporais – por exemplo, que um ferimento no corpo seja acompanhado por dor. Mas o corpo não afeta a mente, nem a mente o corpo; há só a aparência de interação, não a realidade. A Figura 2.1 claramente ilustra a teoria. No entanto, como Deus também é concebido como uma alma, como se supõe que ele estableça as correlações em primeiro lugar? Não servirá responder suavemente que Deus é onipotente e pode fazer qualquer coisa, o problema de como mesmo uma substância divina não-extensa poderia de alguma forma ter um efeito sobre o mundo extenso não-físico permanece.

O ocasionalismo também requer a intervenção de Deus e, consequentemente, está aberto à mesma objeção que o paralelismo. Mas ele é uma teoria menos elegante, na medida em que Deus é representado como tendo de estar presente em toda e qualquer ocasião para assegurar que a correspondência correta ocorra entre os eventos mentais e físicos. Ele está assim condenado a ser um eterno introvertido, diferentemente do Deus de Leibniz, que meramente dá corda e sincroniza os relógios mentais e físicos de uma vez, e depois se retira, deixando que as coisas se ocupem de si mesmas.

### Outros argumentos de Descartes em favor do dualismo

Um argumento de apoio ao dualismo já foi discutido e rejeitado, qual seja, que a capacidade de usar e compreender a linguagem não pode ser explicada em termos mecânicos ou científicos, requerendo, por outro lado, as atividades paramecânicas fantasmáticas de uma alma não-corpórea. No entanto, Descartes tem três outros argumentos principais, que serão examinados nesta seção. Os argumentos são:

1. O argumento da dúvida.
2. O argumento das percepções claras e distintas.
3. O argumento da divisibilidade.

A força dos dois primeiros argumentos é algo parecido com isto: posso pôr em dúvida, ou de algum modo eliminar a lembrança da existência de todos os meus traços físicos. Esses traços assim poderiam, em princípio, serem separados de mim, e não podem, portanto, serem essenciais ao que eu sou. Portanto, eu posso genuinamente existir sem eles. Mas eu não posso do mesmo modo questionar o fato de que eu estou pensando e consciente, porque assim que eu me proponho a fazer isso a própria tentativa será autoanuladora, uma vez que questionar a minha própria existência pressupõe que eu exista e esteja consciente para me engajar nesta atividade.

### O argumento da dúvida

O primeiro argumento, que pode ser chamado de argumento da dúvida sistemática, é facilmente rejeitado, como brevemente veremos. Este argumento resulta do programa de Descartes de tentar questionar tudo, em uma tentativa de descobrir o que é certo e, portanto, pode servir como fundações do conhecimento, sólidas como a pedra. No começo das *Meditations*, Descartes resolve rejeitar não só todas as crenças anteriores que são falsas, mas também qualquer uma que admita a menor dúvida. Mas de quais crenças se pode duvidar? Descartes reconhece que seria uma tarefa sem fim examinar todas as suas crenças uma por uma, e propõe, em vez disso, exa-

Primeiramente → depois

|    |    |    |    |  |
|----|----|----|----|--|
| M1 | M2 | M3 | M4 | (p.ex., querer digitar estas palavras)     |
|    |    |    |    |  |
| F1 | F2 | F3 | F4 | (p. ex., a ação de digitar estas palavras) |

Os eventos mentais (M) correspondem aos eventos físicos (F), mas não há ligação causal alguma entre eles.

Figura 2.1 O paralelismo psicofísico de Leibniz.

minar as fundações do que ele tem tomado como conhecimento. Caso se descubra que essas fundações são falhas, então sua rejeição significará o abandono de todas as crenças que dependem delas.

Descartes inicia salientando que os sentidos algumas vezes o enganaram no que diz respeito a objetos distantes e precisos, e, consequentemente, ele resolve nunca mais colocar total confiança na experiência sensorial. No entanto, baseando-se em reflexões, ele observa que certamente há muitas crenças derivadas dos sentidos das quais não se pode duvidar, tais como a de que ele está agora sentado em frente de um fogo, vestindo um roupão. Mas ele não sonhou que estava sentado desse modo, quando, de fato, ele estava dormindo profundamente na cama? Não é possível que a vida seja nada mais do que um gigantesco e sistemático sonho? Note que, caso esse fosse o caso, você não poderia sequer estar certo de que você tenha um corpo; o tronco e os membros que aparentemente você vê e sente poderiam ser, até onde você possa ver, uma mera invenção da imaginação. No entanto, Descartes reflete que, seja acordado ou dormindo, algumas coisas, tais como as verdades lógicas de que  $2 + 3 = 5$  ou de que um quadrado tem quatro lados, se mantêm, assim que ao menos tais coisas podem ser conhecidas com certeza. Mas não podemos estar enganados mesmo a respeito dessas coisas? Não poderia ser, Descartes supõe, que há um demônio\* infinitamente poderoso e astuto, que o leva ao engano mesmo sobre verdades lógicas? Descartes não acredita literalmente na existência de tal espírito maligno: o demônio pode ser visto como um recurso de imaginação para dar força e capacidade de persuasão às dúvidas céticas de Descartes.

Descartes tece considerações sobre a possível existência de um Deus, cuja existência seja impossível de duvidar e que causa nele a existência dos pensamentos que agora tem. No entanto, ele rapidamente acrescenta, é possível que seus pensamentos, incluindo a ideia de Deus, sejam, nas palavras de Hume, produzidas pela própria energia da mente – em outras palavras, é meramente Descartes, ele mesmo, suscitando seus pensamentos, e não um ser divino que existe separadamente. Mas, se assim for, então ao menos se segue que ele, Descartes, deve ser alguma coisa. Agora não pode ser o caso de ele ser identificado com um corpo humano, porque previamente ele tinha posto em dúvida a existência de um mundo exterior, incluindo seus próprios membros e corpo. Mas caso fosse possível que, até onde ele possa ver, ele não tenha corpo algum, e que sua existência seja meramente uma ilusão, não deveria ele ser persuadido de que ele não existe? No entanto, para ser persuadido de que ele não existe, ele deve em primeiro lugar existir, para que possa, assim, ser persuadido! Desta forma Descartes chega tentativamente à primeira formulação de seu famoso princípio *cogito ergo sum*, “Eu penso, portanto eu existo”. Mesmo se há um gênio maligno usando todo o seu poder para levar enganosamente Descartes a acreditar que há um mundo exterior, incluindo seu próprio corpo, quando não há tal coisa, há ao menos uma coisa sobre a qual ele não pode enganar Descartes, qual seja, a sua própria existência, porque para que Descartes possa ser enganado, ele deve existir:

Sem dúvida, então, eu existo, uma vez que eu sou iludido, e deixe-o iludir-me como ele possa, ele nunca poderá me persuadir de que não sou nada, enquanto eu estiver consciente de que sou alguma

\*N. de T.: No original *demon*, no contexto cartesiano frequentemente traduzido por “gênio maligno”.

coisa. Para que ele deva, em conclusão, ser mantido, todas as coisas sendo madura e cuidadosamente consideradas, que esta proposição Eu sou, Eu existo, é necessariamente verdadeira a cada vez que ela é expressa por mim, ou concebida em minha mente.<sup>19</sup>

É verdade que se Descartes está pensando, então ele deve existir para assim pensar. Mas é importante que fique claro exatamente o que pode e o que não pode ser estabelecido pela reflexão sobre está verdade. Para começar, não é necessariamente verdade que Descartes existe, uma vez que ele poderia não ter existido. Nem mesmo se há tal pessoa como Descartes, isto não garante que ele esteja pensando, então ele deve existir para estar pensando. Mas, para que qualquer coisa possa estar em qualquer estado que seja, isto quer dizer que a coisa em questão deve existir em primeiro lugar. Ao afirmar que Descartes está pensando, a existência do sujeito desta afirmação é pressuposto, assim que continuar a dizer que se Descartes está pensando, então segue-se que ele existe, resulta em um tipo de repetição ou tautologia. No entanto, verdades lógicas como as desse tipo não nos dizem nada a respeito do mundo, e, *a fortiori*, neste caso particular, nada a respeito da natureza do que está fazendo o ato de pensar. Como A. J. Ayer (1910-1989) observou:

Se eu começo com o fato de que eu estou duvidando, eu posso validamente tirar a conclusão de que eu penso, e de que eu existo. Isto quer dizer, se há uma pessoa tal como eu mesmo, então há uma pessoa tal como eu mesmo, e se eu penso, eu penso. Tampouco isto se aplica somente a mim. É obviamente verdade para qualquer um que, se ele existe, ele existe, e se ele pensa, ele pensa. O que Descartes pensou que tinha mostrado foi que afirmações de que ele estava consciente, e de que ele existia, eram, de alguma forma,

privilegiadas, que, para ele ao menos, elas eram evidentemente verdadeiras de um modo que as distinguia de quaisquer outras afirmações de fato. Seu argumento não prova que ele ou qualquer um saiba qualquer coisa. Ele simplesmente estabelece o ponto lógico de que um tipo de afirmação decorre de outra.<sup>20</sup>

Nem reflexões sobre as peculiaridades de certos pensamentos de primeira pessoa, ou elocuções podem estabelecer qualquer coisa sobre a natureza da pessoa que dá vazão a eles. É verdade que se eu consigo proferir ou pensar, "Eu existo", então claramente eu devo existir para ter este pensamento, assim seria incoerente para mim, então, supor que havendo grau proferir ou pensar "Eu existo", eu pudesse vir a ser falso ao mesmo tempo em que eu o expresso publicamente aos outros, ou privadamente para mim. E, de modo inverso, se eu profiro ou penso, "Eu não existo", então simplesmente esta afirmação deve ser falsa, uma vez que a própria condição de ela estar sendo feita por mim é que eu estou lá para fazê-la. No entanto, essas características de afirmações de primeira pessoa, através das quais se elas logram ser feitas significa que elas devem ser verdadeiras na ocasião em que são feitas, e que devem ser falsas se negadas ao mesmo tempo pela pessoa que a faz, não podem possivelmente estabelecer coisa alguma sobre a natureza do agente que as produz. O estudante que responde "Ausente" quando a chamada é feita, ou as pessoas que respondem "Sim" às questões "Você está dormindo?" ou "Você está morto?" devem estar dizendo algo falso, mas isso, obviamente, não estabelece que as pessoas em questão devam ser substâncias não-físicas.

Havendo supostamente estabelecido a verdade de "Eu penso, logo existo" como o primeiro princípio da filosofia que

ele estava buscando, Descartes foi adiante e disse:

Então eu considerei atentamente o que eu era; e eu vi que, embora eu pudesse fingir que não tinha corpo, que não havia mundo, e que não havia lugar algum no qual pudesse estar, eu não podia fingir que eu não existia; ao contrário, a partir do mero fato de que eu pensei duvidar de outras verdades, segue-se evidente e certamente que eu existia... A partir disto eu reconheci que era uma substância, cuja essência total ou natureza é a de ser consciente, e cujo ser não requer lugar algum e não depende de qualquer coisa material. Consequentemente, este eu, quer dizer, a alma por meio da qual eu existo, é inteiramente distinta do meu corpo, e é até mesmo mais facilmente conhecida; e até mesmo se o corpo não existisse de todo, a alma seria simplesmente o que ela é.<sup>21</sup>

Este argumento pode ser colocado formalmente do seguinte modo:

1. Eu posso duvidar de que meu corpo exista.
2. Eu não posso duvidar de que eu exista.
3. Portanto, eu devo ser diferente e distinto do meu corpo.

Tendo supostamente estabelecido a conclusão (3), Descartes então argumenta, em primeiro lugar, que se ele é diferente e distinto de seu corpo, ele deve ser de natureza não-corpórea, isto é, algo desprovido de extensão, e consequentemente de todos os traços físicos; em segundo lugar, porque ele não pode pôr em dúvida ou negar que ele está consciente, a única propriedade verdadeira que inseparavelmente lhe pertence como uma substância não-corpórea é o pensamento ou a consciência.

No entanto, pode o argumento estabelecer em primeiro lugar que Descartes é uma entidade distinta de seu corpo? O

argumento repousa em um princípio conhecido como a Lei de Leibniz. De acordo com a Lei de Leibniz, se uma coisa A e uma B são uma e a mesma coisa, então todas as propriedades de A devem ser as mesmas que todas as propriedades de B. Uma consequência desse princípio é que se há ao menos uma propriedade ou traço que A e B **não** compartilham, então A e B devem ser numericamente coisas diferentes. Por exemplo, se é verdade que o edifício no qual eu agora me encontro fica em Oxford, e se é também verdade que a St Paul's Cathedral fica em Londres, então eu não posso me encontrar na St Paul's Cathedral.

No argumento apresentado para a distinção entre mente e corpo, Descartes supostamente estabeleceu que é possível duvidar de que se tem um corpo, porque a existência do corpo de alguém poderia ser uma ilusão. Mas, ao tentar duvidar de que alguém exista, este alguém deve existir a fim de duvidar, o que quer dizer que tal dúvida não é possível. Portanto, temos algo que é verdadeiro no que diz respeito ao seu corpo, qual seja, que sua existência pode ser posta em dúvida, o que não é verdadeiro de "Eu", si mesmo. Uma vez que o corpo de alguém e "Eu" diferem sob este aspecto, ou seja, que a existência do corpo pode ser posta em dúvida, mas a existência do "Eu" não pode, então, de acordo com a Lei de Leibniz, parece resultar que a pessoa ou consciência não pode ser idêntica ao corpo, mas deve ser diferente e distinta dele.

Antes de continuar com esta discussão, veja o Exercício 2.7.

### EXERCÍCIO 2.7

O argumento de Descartes a partir da dúvida sistemática tem êxito? Você pode construir um argumento paralelo, em que embora uma coisa A possa ser posta em dúvida, e outra coisa B não possa ser

posta em dúvida, no entanto está claro que A e B são uma e a mesma coisa? Se tal argumento pode ser produzido, ele pode ser usado para revelar um vício de forma no pensamento de Descartes. Faça isso sozinho e depois discuta com alguém mais, antes de prosseguir na leitura.

Há certas exceções à Lei de Leibniz. Neste caso deve ser reconhecido que argumentos que empregam verbos mentais intencionais, como “crer”, “duvidar”, “imaginar”, “sonhar” e assim por diante em sua construção, não podem ter êxito no estabelecimento de não-identidade de itens. Estes estados mentais, lembrem-se (veja Capítulo 1, seção “Sobre a ‘sobredade’”), mostram a característica de “sobredade” (*aboutness*), isto é, eles possuem um conteúdo representacional. Além disso, este conteúdo representacional é **aspectual**, uma vez que um e o mesmo estado de coisas pode ser representado sob diferentes aspectos. Quando Édipo casou-se com sua mãe, ele não se deu conta de que ela era sua mãe. Ele a conhecia somente como a Rainha de Tebas. Do mesmo modo, para adaptar um exemplo de Daniel Dennett, eu poderia ser conhecido de todos em conjunto como um inocuo professor de filosofia, e não como o Estrangulador Maluco. No entanto, eu ainda poderia vir a ser o Estrangulador Maluco. Tendo essas observações em mente, é fácil formular um argumento paralelo ao de Descartes, que claramente falha. Veja o argumento:

1. Lois Lane sabe que o repórter com o qual ela trabalha é Clark Kent.
2. Ela não sabe que o repórter com o qual trabalha é o Super-Homem.
3. Portanto, Clark Kent não pode ser o Super-Homem.

Claramente alguma coisa deu errada aqui, porque um aspecto integral das

histórias de Super-Homem é que o Super-Homem é uma e a mesma pessoa que o superficialmente inapto e não assumido Clark Kent. Por que a Lei de Leibniz não se sustenta pode ser facilmente visto. Só porque alguém – no caso Super-Homem/Clark Kent – é conhecido por Lois sob uma descrição, não implica que ela o conhecerá sob descrições alternativas. Ao conhecer o homem de aço sob o aspecto de Super-Homem, não se segue que Lois Lane reconhecerá este mesmo homem quando ele retira sua capa e coloca seus óculos de coruja. No entanto, o fato de que ela, erradamente, crê que Clark Kent não é o Super-Homem não pode estabelecer que Clark Kent e o Super-Homem não sejam um e o mesmo indivíduo.

Por meio deste padrão de argumento, então, Descartes não pode estabelecer que ele é uma substância não-corpórea, distinta de seu corpo. Porque, no entanto, ele acredita ter estabelecido isto, e consequentemente logrado desvencilhar-se de todos os seus traços corporais, deixando para trás o único atributo do pensamento ou consciência, ele conclui não meramente que ele é uma coisa pensante – uma conclusão que poderia ser cabível –, mas que ele é **somente** uma coisa pensante, isto é, uma coisa cuja única essência ou natureza consiste em **nada mais** do que consciência. Uma vez que agora estou consciente, enquanto escrevo isto, segue-se que sou uma entidade consciente, isto é, a consciência é um traço meu. No entanto, não se segue que **tudo** o que eu sou seja consciência, que este traço esgotte totalmente a minha natureza. Poderia ser que para pensar que eu tenho que ser corporificado e que certos complexos processos físicos devem continuar dentro de mim. Hobbes, um contemporâneo de Descartes, fez a crítica essencial a este argumento quando escreveu: “Pode ser que a coisa que está consciente seja o sujeito de

uma mente, razão ou intelecto, e, assim, pode ser algo corpóreo; o contrário, assume-se, não está provado".<sup>22</sup> O argumento de Descartes a partir da dúvida não pode se mostrar de outra forma.

### **O argumento a partir da percepção clara e distinta**

Na Sexta Meditação, no entanto, Descartes apresenta outro argumento para estabelecer a distinção entre a mente e o corpo e a possibilidade da separação deles, que é muito mais difícil de suprimir. O argumento apresenta-se da seguinte forma:

Eu sei que qualquer coisa que eu comprehendo clara e distintamente pode ser feita por Deus assim como eu o comprehendo; então a minha habilidade de compreender uma coisa clara e distintamente separadamente uma da outra é suficiente para me assegurar que elas são distintas, porque Deus ao menos as pode separar... Agora, eu sei que existo, e, ao mesmo tempo, observo que não há absolutamente nada que pertença à minha natureza ou essência exceto o mero fato de que eu sou um ser consciente; a partir somente disto posso inferir validamente que minha essência consiste simplesmente no fato de que eu sou um ser consciente. É realmente possível [...] que eu tenha um corpo estreitamente ligado a mim mesmo, mas ao mesmo tempo eu tenho, por outro lado, uma ideia clara e distinta de mim próprio tomada simplesmente como uma consciência, não um ser extenso, e, por outro lado, uma ideia distinta de corpo, tomada simplesmente como um ser extenso, não consciente; consequentemente, é certo que sou distinto de meu corpo, e poderia existir sem ele.<sup>23</sup>

Este argumento requer um comentário mais longo. Como Anthony Kenny salienta, naquele tempo os críticos assinalaram que

Não se segue do fato de que a mente humana pensando sobre si própria não se perceber como sendo outra coisa senão uma coisa que pensa, que sua natureza ou essência consista somente em ser uma coisa que pensa, no sentido de que esta palavra **somente** exclui todas as outras coisas que também poderiam pertencer supostamente à natureza da alma.<sup>24</sup>

Isto é uma reminiscência do comentário de Hobbes citado anteriormente. Kenny também destaca que é fatalmente fácil confundir as duas afirmações seguintes:

1. Eu não constato coisa alguma que pertença à minha natureza ou essência, à exceção do pensamento.
2. Eu constato que nada mais realmente pertence à minha natureza ou essência, à exceção do pensamento.

O fato de que Descartes não logra perceber que alguma coisa mais pertença à sua natureza ou essência à exceção do pensamento não significa que nada mais assim pertença. A afirmação 2 não decorre da afirmação 1, mas é da afirmação 2 que Descartes necessita para estabelecer que ele seja uma coisa cuja essência em sua totalidade consista em pensar ou ter consciência simplesmente, compartilhando nenhum traço físico que seja, e, consequentemente, estabelecendo que ele possa realmente existir separado e isoladamente de seu corpo.

Será útil fazer a colocação formal do argumento, antes de proceder à sua análise mais aprofundada; uma formulação provisória pode ser a seguinte:

1. Eu tenho uma compreensão clara e distinta de mim próprio como uma coisa consciente, mas não extensa.

2. Eu tenho uma consciência clara e distinta de meu corpo como uma coisa extensa, mas não consciente.
3. Portanto, eu posso clara e distintamente compreender a mim próprio como uma coisa não-extensa e consciente, separada de meu corpo extenso, e eu posso clara e distintamente compreender meu corpo como uma coisa extensa e não-consciente separada de mim próprio como uma coisa consciente.
4. O que quer que eu possa clara e distintamente compreender pode ser realizado por Deus (como eu o comprehendo)
5. Portanto, Deus pode fazer com que eu, como uma coisa consciente, não-extensa possa existir separadamente de meu corpo extenso, não-consciente.
6. Se Deus pode fazer com que eu e meu corpo possamos existir separadamente, então eu e meu corpo realmente podemos existir separadamente.
7. Consequentemente, eu sou distinto de meu corpo e posso existir separadamente dele. Eu posso continuar a existir como uma consciência não-física independentemente de todos os meus traços corpóreos.
8. Consequentemente, meus traços corpóreos não podem ser essências à minha existência.
9. Portanto, eu devo ser uma coisa completamente não-corpórea, que é essencialmente uma coisa consciente, não-extensa, uma *res cogitans*.

Após haver estudado este argumento, veja o Exercício 2.8, antes de prosseguir sua leitura.

### EXERCÍCIO 2.8

Você pode ver algum problema com o argumento apresentado? É difícil avaliá-lo, por isso sugiro que você examine a premissa 3 cuidadosamente. Ela é ambígua e pode ser interpretada com dois significados, um dos quais dará a Descartes o que ele quer, desde que seja verdadeira (o que pode não ser)..

Como eu sugeri no Exercício 2.8, um problema com o argumento é o de que a terceira premissa é ambígua. Como M. D. Wilson sugere, Descartes poderia estar afirmando:<sup>25</sup>

- a. Eu posso distintamente conceber minha mente e meu corpo separadamente um do outro.
- b. Eu posso distintamente conceber minha mente e meu corpo **como sendo separados** um do outro.

Descartes não vê a diferença entre (a) e (b) e poderia estar afirmando qualquer um dos dois; ele necessita estabelecer a verdade de (b), não a de (a). (a) poderia ser verdade, independentemente de (b), mas (a) é fraca demais e não lhe levará à conclusão que ele deseja, isto é, que a mente pode existir independentemente do corpo. Para avaliar isto, veja a seguinte expansão de (a):

- a1. Você pode fazer uma concepção de minha mente, isto é, pensar ou refletir sobre meus traços mentais, sem ao mesmo tempo pensar sobre meus traços físicos. Por exemplo, eu poderia descrever o meu *eu* como emocionalmente frágil, estúpido, de compreensão lenta e imperceptível. Todos esses traços são psicológicos, que não fazem menção a minhas caracte-

rísticas físicas. Do mesmo modo, eu poderia mencionar muitas de minhas características físicas – baixo, barbudo, de olhos azuis, gordo – sem dizer coisa alguma sobre meu lado mental. Neste sentido, as duas concepções podem ser formuladas independentemente uma da outra. Mas isto não pode mostrar que eu possa existir mentalmente separado de meus traços físicos.

Para verificar isto, compare a extensão e a forma de um corpo. Eu posso falar sobre a extensão sem fazer referência ou menção explícita à sua forma, ou falar de sua forma sem mencionar sua extensão explicitamente. Neste sentido, a forma e a extensão podem ser pensadas ou concebidas independentemente uma da outra. Porém, de forma alguma se segue que os objetos de tais concepções possam ser concebidos existir separadamente um do outro. Logo nos damos conta disto quando lembramos que o que é extenso deve ter uma ou outra forma, embora mutável, e o que possui uma forma deve possuir extensão. A noção de uma coisa que tenha forma, mas careça de extensão, e vice-versa, é incoerente. No entanto, a impossibilidade lógica de haver forma sem extensão, ou extensão sem forma, não nos impede necessariamente de falar sobre extensão sem fazer qualquer menção explícita de forma, ou de falar sobre forma sem fazer qualquer menção explícita de extensão.

Para que Descartes tenha alguma chance sequer de estabelecer o dualismo por meio deste argumento, o que ele necessita para atuar como premissa 3 não é meramente que ele possa clara e distintamente conceber traços mentais separadamente de traços físicos – isto é, cada **concepção** pode ser mantida separada-

mente, ou independentemente, da outra – mas, melhor, que os **objetos** dessas concepções, sobre o que são essas concepções – neste caso, a mente e o corpo, em vez das próprias concepções – possam ser concebidas existir separadas e independentemente uma da outra. Em outras palavras, em vez de (a1), o que Descartes necessita é de:

b2. Eu posso clara e distintamente conceber minha mente como existindo independentemente e separada de meu corpo.

Esta afirmação foi atacada por um dos contemporâneos de Descartes, o filósofo Antoine Arnauld (1612-1694). A essência da objeção de Arnauld era que Descartes poderia acreditar que ele pudesse conceber clara e distintamente a si mesmo como um puro sujeito de consciência capaz de existir separadamente de seu corpo físico, mas na realidade ele não pode formar esta concepção. Suponha que Descartes, em vez disso, tivesse argumentando que ele pudesse clara e distintamente conceber que um triângulo euclidiano tivesse a propriedade de ser de ângulo reto, sem, portanto, clara e distintamente conceber que ele obedecesse ao teorema de Pitágoras, e que, portanto, seria possível existirem triângulos de ângulo reto aos quais o teorema de Pitágoras não se aplicasse. Esta conclusão é inaceitável, uma vez que sabemos, e pode ser facilmente demonstrado, que a propriedade de ser um triângulo de ângulo reto e a propriedade de satisfazer ao teorema de Pitágoras estão logicamente conectadas. Em um caso como este, Descartes estaria forçado a concluir que ele não poderia clara e distintamente conceber o que ele originalmente pensou que poderia. O que parece ser concebível à primeira vista, após uma análise mais aprofundada, mostra-se ser

de forma alguma concebível, e, portanto, impossível.

Descartes não se impressionou com a objeção de Arnauld. Em sua resposta, ele introduziu e faz uso da distinção entre seres completos e incompletos. Descartes nos solicita tomar em consideração um corpo em movimento com uma determinada figura ou forma. A distinção entre forma e figura é formal, e é possível formar uma concepção de movimento sem falar de figura e vice-versa. Do mesmo modo, é possível formar a concepção de um triângulo sem pensar que ele obedeça ao teorema de Pitágoras. Por outro lado, não é possível formar uma concepção completa de movimento separadamente de uma coisa que está se movendo, e, do mesmo modo, não pode haver uma concepção completa de um triângulo ao qual o teorema de Pitágoras não se aplique. Portanto, o movimento pode haver separadamente de algo que se move, e um triângulo de ângulo reto que não exemplifique o teorema de Pitágoras não pode existir.

No entanto, no caso do corpo e da mente nada disto se aplica, sustenta Descartes, porque é impossível formar concepções do corpo e da alma como seres completos, ou, ao menos, concepções do corpo e da alma que, se não **totalmente** completas, sejam ao menos **sufficientemente** completas, para garantir a possibilidade que uma exista separadamente da outra.

Descartes escreve:

Mas eu comprehendo completamente o que um corpo é, meramente pensando nele como extenso, provido de figura, móvel, etc. e negando a ele todas aquelas coisas que pertencem à mente; e vice-versa, eu comprehendo a mente como uma coisa completa, que duvida, comprehende, deseja e assim por diante, embora eu negue que qualquer uma das coisas contidas na ideia de corpo esteja nela.<sup>26</sup>

Em outras palavras, o que Descartes está afirmando é que posso formar um conceito distinto e suficientemente claro e completo de mim mesmo como uma coisa pensante sem extensão, e, igualmente, posso formar um conceito distinto e suficientemente claro e completo de meu corpo como uma coisa extensa sem pensamento. Os objetos dessas concepções – isto é, que as concepções são sobre, ou dizem respeito a, como opostas às próprias concepções – realmente parecem ser distintos e suficientemente completos em si mesmos. Na concepção de processos físicos, estou pensando em eventos que são localizados espacialmente e ocupam uma área ou volume de espaço. No caso da atividade cerebral, escaneamentos do cérebro podem mostrar fluxos sanguíneos e atividades neurológicas ocorrendo em áreas específicas do cérebro. A forma e o volume da área do cérebro envolvida em atividade neural podem ser mensurados, assim como a frequência dos impulsos nervosos através dos nervos. Essa caracterização do que se passa no cérebro parece ser suficientemente completa para que o estado de coisas exista exatamente como é descrito, sem posteriores adições a ele. Por outro lado, parece que somos capazes de alcançar uma concepção suficientemente clara da natureza dos traços mentais que existem por si mesmos, em completa independência dos traços físicos.

No entanto, minha mente e meu corpo podem ser concebidos como coisas completas em si mesmas, isto é, não incluindo ou tomando emprestado traço algum um do outro, então é possível que estes estados de coisas possam existir separadamente um do outro; e uma vez que Deus pode criar aquilo que eu posso conceber clara e *disjuntivamente* como possível, minha mente é meu corpo realmente podem existir separadamente um do outro.

Assim Descartes escreve:

A mente pode ser percebida clara e distintamente, ou suficientemente; assim, para que seja considerada uma coisa completa, sem qualquer uma daquelas formas ou atributos por meio dos quais reconhecemos que o corpo é uma substância ... um corpo é compreendido distintamente como uma coisa completa sem aqueles [traços] que pertencem à mente.<sup>27</sup>

Pode ser feita a objeção de que isto não leva em conta a crítica de que, ao verificar que tudo que parece necessário para que alguém exista são suas experiências conscientes e percepção em geral, isto não leva em conta aquilo que não logramos observar, e poderia ser que estes traços não percebidos sejam físicos e igualmente essenciais à existência de alguém. Mas Descartes tem pronta a sua resposta.

Certamente, quando provei que Deus existe ... quem pode fazer tudo o que eu clara e distintamente conheço ser possível. Pois embora muito em mim exista que eu ainda não observe [no estágio particular das meditações que tinha alcançado]... no entanto, uma vez que aquilo que eu efectivamente noto é suficiente para que eu subsista somente com isto [sem outras palavras, sua concepção de si mesmo como um ser consciente], eu estou certo de que eu poderia ter sido criado por Deus sem outros [atributos] que eu não observo.<sup>28</sup>

Não será suficiente responder ao argumento de Descartes a partir da percepção clara e distinta de que o máximo que ele pode mostrar é que é **possível** para a mente e o corpo existirem separadamente, e que ele não mostra que a mente e o corpo **realmente** existam independentemente um do outro. Isto porque tudo o que Descartes necessita para refutar o materialismo é a pura possibilidade de a mente e o corpo serem separados segundo o modo que ele propõe. Caso seja genuinamente concebível que estados mentais devam

existir independentemente de estados físicos, então a fisicalidade não poderia ser a própria essência de estados mentais, o *sine qua non* da existência deles, e todas as tentativas de reduzir a mente à matéria devem, em última instância, falhar.

No entanto, caso se resista à conclusão de Descartes de que a mente possa existir separadamente do corpo, então parece que devemos novamente lançar mão de uma observação simplificada, qual seja, a de que, contrariamente ao que Descartes imagina, ele não pode genuinamente formar uma concepção distinta e suficientemente clara de si mesmo como uma coisa puramente pensante capaz de existir separadamente de um corpo físico. Ele pode, aparentemente, alcançar uma concepção clara e distinta de si mesmo como um sujeito de consciência não-físico, mas somente porque ele não tem conhecimento de sua verdadeira natureza.

Ao afirmar que não somos nada além de mentes não-físicas, Descartes assume que o modo como a nossa experiência consciente nos é apresentada capta suficientemente a natureza dela para garantir que podemos realmente existir independentemente de todos os traços físicos. Mas isto é realmente assim? Em algum lugar no tempo a natureza do som não era transparente para nós. Ninguém sabia que o som consistia de ondas e que estas requeriam um meio no qual pudessem ser propagadas. Somente após a realização de experimentos, tais como evacuar o ar de dentro de um recipiente no qual uma campainha estava soando e observando que o som gradualmente diminuía e finalmente cessava de existir, quando se obtinha um vácuo, que foi percebido que é da própria natureza do som consistir de ondas em um meio tal como o ar. Do mesmo modo, embora presentemente não tenhamos a menor ideia de como a consciência poderia ser idêntica ou depender de pro-

cessos cerebrais para sua existência, não pode ser descartado antecipadamente que seja da própria natureza da consciência ter processos físicos em seu âmago. Entidades conscientes, onde quer que existam no universo, poderiam ser necessariamente arranjos físicos de um ou outro tipo, embora atualmente careçamos totalmente de algum conjunto plausível de conceitos ou de teorias que possam ser empregadas para tornar transparente a identidade do mental com o físico, ou, alternativamente, a dependência dele sobre o físico.

Em seu livro sobre Descartes, John Cottingham dá uma resposta semelhante:

Parece plausível argumentar que, embora o conceito de pensamento seja bastante distinto do conceito de atividade cerebral, o pensamento é, não obstante, um processo funcional, que não pode operar sem algum tipo de *hardware* (seja um cérebro ou algo análogo). Engenheiros de *softwares*, para terem certeza, elaboram seus programas em termos puramente abstratos, sem qualquer referência ao mundo físico; mas eles sabem que para que seus programas possam operar realmente eles devem ser fisicamente incorporados (por exemplo, em um disco rígido). Pois haver um programa de *software* operacional na ausência de um substrato físico é, em última instância, uma noção incoerente: não é somente que ela não ocorra em nosso universo; não há mundo possível algum no qual seja encontrada (assim como não há um mundo possível no qual exista um processo digestivo na ausência de algum tipo de órgão físico capaz de fazer tal tarefa). Se isso está correto, então, embora possa parecer plausível à primeira vista supor que seja logicamente possível a existência de mentes separadas de corpos, a noção torna-se, em última instância, incoerente, e assim falha o argumento de Descartes.<sup>29</sup>

No entanto, o argumento de Cottingham assume a verdade do funciona-

lismo, a visão segundo a qual a mente é como um programa rodado no *hardware* do cérebro, e, como veremos no Capítulo 5, o funcionalismo é profundamente falho como uma teoria da mente. Mais contrariamente, o funcionalismo não pode por si só estabelecer a verdade do materialismo, uma vez que funções, como os funcionalistas prontamente admitiram, deveriam ser pensadas não como fantasmáticas e mentais, nem como processos físicos, mas como abstrações que são neutras entre as duas. As funções, foi dito, poderiam ser instanciadas no *hardware* etéreo de substâncias imateriais e não somente em conjuntos de arranjos materiais. Assim, pode haver mundos possíveis nos quais, mesmo que o funcionalismo viesse a dar um retrato da mentalidade verdadeiro, mas incompleto, as mentes terminam sendo, ao final de contas, imateriais. Mas, retomando o ponto anterior, caso seja possível que mentes algumas vezes existam independentemente de arranjos, um materialismo profundamente reducionista será falso.

### O argumento da divisibilidade

Este argumento pode ser breve e formalmente exposto da seguinte forma:

1. O corpo é divisível em partes.
2. A mente não é divisível em partes.
3. Portanto, a mente deve ser de uma natureza inteiramente diferente da natureza do corpo, isto é, ela deve ser essencialmente não-física.

Quando Descartes considera seus próprios estados mentais – por exemplo, quando ele está com uma dor, ou bravo, ou pensado sobre o que ele comerá no jantar –, ele diz não poder distinguir nenhuma parte dentro de si próprio. Mas isso não deveria se constituir em surpresa. Consideradas

como **estados** de um sujeito, as experiências não podem ser separadas em metades ou quartas partes, porque não faz sentido falar de estados, sejam eles físicos ou mentais, como compostos de partes. Considere um exemplo não-mental, um componente químico, digamos, nitroglicerina, que está em um estado estável. O que isso quer dizer é que sua condição é tal que não há perigo de que ela exploda sem aviso. A nitroglycerina como tal pode ser dividida – várias porções poderiam ser colocadas em diversos recipientes –, mas o que possivelmente poderia significar descrever seu estado de estabilidade como sendo divisível? Estados não são divisíveis, mas não pela razão de serem coisas peculiares, simples, indivisíveis, mas, sim, porque não são coisas, tampouco substâncias lógicas, de todo. (Devemos nos preaver de um erro, do qual já preveni anteriormente, o engano de interpretar todos os substantivos como se eles adquirissem seus significados por meio da nomeação de substâncias. Veja seção “Dualismo substancial”). O mesmo também é verdadeiro quanto às capacidades. Nitroglycerina tem a capacidade de explodir, mas qual seria o sentido de falar de compartilhar porções desta capacidade, como opostas à própria nitroglycerina? Do mesmo modo, se, por exemplo, uma pessoa possui força, ou a habilidade de correr muito rapidamente, não podemos falar de cortar pela metade sua habilidade de correr ou a capacidade de levantar objetos pesados, assim como poderíamos falar de dividir um pedaço de papel em dois, usando uma tesoura ou um cortador de papel. Certamente, a habilidade de uma pessoa para levantar objetos pesados ou cobrir distâncias com rapidez pode declinar, ou até mesmo desaparecer, mas isso não quer dizer que ela parou de existir porque literalmente fragmentou-se como uma noz se estilhaça quando damos nela uma martelada.

Não deveríamos dizer “Corpos são divisíveis, mentes não são” porque, tomando emprestada uma expressão do filósofo Ludwig Wittgenstein, isso faria com que a diferença entre corpos e mentes parecesse demasiadamente pouca. Wittgenstein nos convida a fazer a comparação da afirmação “Três pessoas não podem se sentar lado a lado neste banco; não tem espaço” com a afirmação “3x18 polegadas não entrarão em 3 pés”. Essas afirmações parecem ser muito similares, mas de fato são muito diferentes. A primeira afirma uma impossibilidade física, a segunda uma impossibilidade lógica, algo que é excluído pelos conceitos de 3x18 polegadas e de 3 pés.

A afirmação “mentes não são divisíveis” se assemelha à segunda afirmação, não à primeira, e marca uma impossibilidade conceitual, algo que não faz sentido dizer a respeito de mentes, assim como não faz sentido perguntar se o número 3 tem pais. O fato de o número 3 não ter pais não faz dele um órfão!

Reiterando: mentes não são divisíveis, mas não porque sejam algum tipo miraculoso de átomo espiritual duro, indivisível, mas porque os correlativos conceitos de divisibilidade e indivisibilidade não têm aplicação alguma a estados e capacidades mentais, devido ao fato de eles não serem substâncias lógicas, coisas capazes de existir por si mesmas.

## Problemas para o dualismo

### **A posse de uma mente é tudo ou nada**

Para o cartesiano, somente os seres humanos têm mentes sob a forma de almas incorporadas. A racionalidade e a consciência residem na alma; portanto, os animais devem ser considerados como

autômatos não-conscientes. No entanto, isto faz com que o contraste entre nós e os demais seres vivos seja demasiadamente rígido. Em vista dos fatos da biologia evolutiva, incluindo, principalmente, o crescimento em tamanho e complexidade do sistema nervoso central à medida que ascendemos na escala evolucionária, é certamente mais plausível considerar a posse de uma mente como um espectro que abarca desde criaturas como formigas e lesmas, que podem ter somente a mais rudimentar das vidas mentais, à completa panóplia de estados mentais que nós próprios possuímos. A ausência de mente é uma questão de grau, e não uma questão de tudo ou nada, como insistem os dualistas.

### **O cérebro é destituído de uma função**

Se a existência de mentes, em última instância, não tem nada a ver com o que ocorre no cérebro e no sistema nervoso em geral, por que temos um órgão tão imensamente complicado, a coisa mais complexa que conhecemos? Por que as pessoas sofrem limitações mentais se o cérebro é danificado ou experiem estados de consciência alterados quando tomam drogas, se o cérebro não é responsável por nossas vidas mentais? Um dualista dedicado pode resistir a essa conclusão, insistindo que o cérebro é meramente semelhante a uma estação de transmissão que transmite as mensagens da alma. Os maus funcionamentos da mente ou da alma são somente aparentes; o que está realmente acontecendo é que o corpo está deixando a mente cair, quer não executando os seus comandos, quer executando-os imperfeitamente. Sob este ponto de vista, a inabilidade de uma pessoa em compreender e produzir a fala quando o centro da fala de seu cérebro foi severamente danificado é

meramente o resultado de o cérebro não mais estar capaz de traduzir as instruções provindas da alma. Puramente, dentro da própria alma, a habilidade de compreender e produzir a linguagem permanece intacta e não perturbada. Esta é precisamente a linha adotada por Descartes em uma resposta a um de seus críticos, Pierre Gassendi:

Pois, a partir do fato de que ela [a alma] não age de forma tão perfeita no corpo de uma criança quanto no de um adulto, e que a ação dela possa seguidamente ser impedida pelo vinho ou outras coisas corporais, segue-se somente que, enquanto a mente for unida ao corpo, ela usa o corpo como um instrumento naquelas operações nas quais ele é usualmente ocupado, não que ela se torne mais perfeita, ou menos, por meio do corpo... Tenho ... com frequência distintamente mostrado que a mente pode operar independentemente do cérebro; pois certamente o cérebro não pode ter serventia alguma para a compreensão pura, mas somente à imaginação e à sensação.<sup>30</sup>

Quanto à plausibilidade dessa resposta, deixo a decisão ao próprio leitor.

### **A falácia do homúnculo**

O termo “homúnculo”, que quer dizer “homem pequeno” ou “anão”, provém dos primeiros biólogos que pensavam que cada espermatozoide humano continha dentro de si um ser humano em miniatura, que entra no ovo, no momento da concepção, e, então, gradualmente, se desenvolve em um feto humano. Os traços essenciais para a constituição de um ser humano tinham de estar presentes e inteiramente localizados desde o princípio, de acordo com os primeiros biólogos, porque não tinham ideia alguma de como um complicado organismo vivo poderia

ser gradualmente formado a partir de materiais que originalmente não eram vivos. Mas se um homúnculo é postulado para explicar como um ser humano em sua inteiridade vem a ser, então a teoria que faz uso de tal recurso é vazia ou circular, porque sua base explicativa contém o próprio fenômeno que ela pretende explicar.

A falácia do homúnculo é evidente na tentativa que Descartes faz de explicar a percepção por meio da manutenção de que a alma volta-se em direção à glândula pineal e inspecciona as imagens projetadas em sua superfície como resultado de impulsos transmitidos pelos órgãos sensoriais externos. Este modelo imediatamente convida à questão de como se supõe que a própria alma perceba. Tem ela também um mecanismo fantasmático em duplicata à glândula pineal em seu interior, por meio do qual ela perceba? Será óbvio que embarcar nesta linha de pensamento conduziria a uma viciosa regressão infinita de agentes de percepção no interior de agentes de percepção, à semelhança de uma série de bonecas russas encastradas, exceto que não pode haver uma última boneca no miolo. Em geral, qualquer teoria psicológica que atribua à alma, ou ao cérebro, aqueles traços ou habilidades específicos aos quais pretendem explicar comete a falácia do homúnculo.

### **O problema de contar almas**

O ataque mais penetrante ao cartesianismo, segundo meu ponto de vista, deriva-se do filósofo alemão Immanuel Kant (1724-1804). Descartes afirma que uma, e somente uma, alma está associada com um dado corpo humano e age por meio dele. Mas, pergunta Kant, o que diria o cartesiano à sugestão de que, tanto quanto ele possa dizer, milhares ou milhões de tais almas, todas pensando o mesmo pensamento, falem por meio da boca de um

ser humano individual? Além disso, ao longo do tempo, por que não deveria uma sucessão de substâncias de almas serem associadas a um ser humano, cada substância de alma transmitindo seus estados ao próximo membro da série, do mesmo modo que cada membro de uma série de bolas elásticas, cada bola sendo batida e batendo por sua vez em outra bola transmite seu *momentum* à próxima bola na fileira, e assim por diante, ao longo da série?

O ponto que Kant está levantando é fundamentalmente este: Como se pode contar almas? Com que critério ou critérios se decide a um momento dado quantas almas estão presentes? Quando se pode dizer que uma alma termina, por assim dizer, e estamos passando à próxima? Em outras palavras, como uma alma é individuada, isto é, dita separada como um indivíduo distinto e separado, de outra alma em um dado momento? Além disso, como se pode dizer que ao longo do tempo almas sejam separadas? Como é que se pode distinguir a situação na qual uma e a mesma alma seja encontrada novamente após um lapso de tempo daquela situação na qual uma alma qualitativamente similar, mas qualitativamente distinta é encontrada?

O ponto de Kant é que, se a experiência não fornece embasamento para distinguir entre o caso em que o conceito de "uma e a mesma alma" é aplicável, e a situação em que o conceito de "alma qualitativamente similar, mas numericamente diferente" pode ser aplicado, então realmente não temos conceito algum do que significamos por uma alma, e, portanto, nenhum conceito de uma coleção de tais/indivíduos. Mas então não temos conceito algum de alma como substância individual, e a noção de substância incorpórea deve ser rejeitada por ser destituída de significado.

Estabeleça o contraste com o caso de bolas de bilhar de uma cor determinada. Uma bola de bilhar se parece muito com outra, e, na prática, pode ser muito difícil estabelecer a diferença. Mas, claramente, não obstante o quanto duas ou mais dessas bolas se pareçam, deve sempre haver um modo fundamental para que elas possam ser distinguidas umas das outras, ou seja, nenhuma bola pode ocupar o espaço que outra ocupa em um e mesmo tempo. E, por extensão a este princípio, nenhuma bola pode ocupar os mesmos espaços sucessivos ocupados por outra bola aos mesmos tempos sucessivos. A rota espacial traçada por uma bola em um dado período de tempo deve ser único àquela bola. O que se aplica a bolas de bilhar aplica-se igualmente a seres humanos considerados como criaturas corpóreas de carne e de sangue. Não obstante o quanto Rosie se pareça com sua irmã gêmea idêntica Alice, ela não pode ocupar o mesmo lugar que Alice ao mesmo tempo, tampouco os mesmos lugares que Alice nos mesmos momentos ao longo de um período de tempo. Vejamos o Exercício 2.9 antes de prosseguir neste assunto.

### EXERCÍCIO 2.9

Você vê algum problema na tentativa de explicar a identidade das almas do mesmo modo que explicamos a identidade de bolas de bilhar e pessoas de carne e sangue? Depois de haver trabalhado sozinho sobre esta questão, discuta-a com outra pessoa, antes de prosseguir a sua leitura.

A dificuldade com a qual nos deparamos neste exercício deve ser bastante óbvia. É que substâncias incorpóreas, de acordo com sua própria natureza, não ocupam uma área especial de espaço em

um tempo dado, e, portanto, este fundamento de identidade não pode ser aplicado. Mas, se isto é assim, não temos como distinguir uma alma da próxima, e disso resulta que temos de admitir que não possuímos um conceito claro e coerente da alma. Kant salienta que o mero fato da própria autoconsciência de alguém não ajuda. Na *Critique of Pure Reason*\* ele escreve: “A identidade da consciência de mim mesmo em tempos diferentes é... somente uma condição formal de meus pensamentos e da coerência deles, e de forma alguma prova a identidade numérica de mim mesmo como sujeito.”<sup>31</sup>

O ponto que Kant está levantando é o seguinte: é necessariamente verdade que, no que concerne a qualquer pensamento ou estado mental que seja meu, é de fato um e o próprio mesmo “eu” que tem essas experiências. (Kant, em uma ocasião, fez esta colocação ao dizer que, no que concerne às aquelas experiências que são minhas, deve ser possível para o “eu penso” acompanhar aquelas experiências.) Mas, embora seja necessariamente verdade que seja um e o mesmo “eu” que esteja consciente de todos os meus pensamentos e experiências como uma condição para que tais pensamentos e experiências sejam **meus** pensamentos e experiências, isto de forma alguma prova que eu, em última instância, seja uma substância lógica incorpórea, não-composta, uma alma cartesiana.

Para poder demonstrar isso, teria de ser capaz de especificar quais características de minha experiência me autorizariam aplicar e usar o conceito de alma. Mas, para fazer isso, a experiência teria de me oferecer um objeto, discernível pelos sentidos, que persista ao longo do tempo. Eu devo, ao menos em teoria, ser capaz de distinguir este objeto de outros objetos semelhantes, em qualquer mo-

\*N. de T.: Em português: *Crítica da Razão Pura*. São Paulo: Ícone, 2007.

mento dado, e ser capaz de reidentificá-lo ao longo do tempo. No entanto, como tentei deixar claro, isso não pode ser feito no caso de almas, nem mesmo no meu próprio caso, uma vez que uma coisa incorpórea é um objeto do qual, por definição, não posso ter experiência, porque o que carece de todas as características físicas – e, em particular, extensão no espaço – não pode, em princípio, ser detectado pelos sentidos. Mas, como Kant e seu antecessor, Hume, concordam, conceitos que não têm aplicação dentro do escopo da experiência são vazios de conteúdo. Portanto, o conceito de uma pessoa como uma alma incorpórea, não-corporificada, pareceria, em última instância, ser sem sentido. Por outro lado, se pensarmos em pessoas como entidades familiares de carne e sangue, que encontramos nas ruas, então podemos ter um conceito genuíno e significativo de uma pessoa. Isso porque, pensadas como coisas corporificadas, podemos, em última instância, sempre distinguir uma pessoa de outra em um momento dado, ou o longo de um período de tempo, uma vez que não podem ocupar o mesmo espaço ao mesmo tempo, ou os mesmos espaços em tempos sucessivos. Com esse critério, posso distinguir Alice de Rosie, mesmo que sejam gêmeas idênticas, e, igualmente, tanto Alice quanto Rosie podem me distinguir de outras pessoas que encontrem.

Este modo de conceber pessoas encontra sua expressão nas palavras de Kant:

Posso ainda assumir que a substância que em relação ao nosso sentido externo possui extensão é em si mesma possessora de pensamentos, e que esses pensamentos podem por meio de seu próprio sentido interior ser conscientemente representados. Deste modo, o que em uma relação é intitulado incorpórea, seria em outra relação ao mesmo tempo um ser pensante, cujos pensamentos não podemos intuir;

embora possamos intuir, de fato, seus sinais no campo da aparência.<sup>32</sup>

Com esta última frase, Kant quer dizer que não podemos observar diretamente os pensamentos de outras pessoas, mas podemos ver os sinais do que as pessoas estão pensando pelos modos como se comportam, e, em particular, por suas expressões faciais. Ele prossegue:

Em consonância, a tese de que as almas (como tipos particulares de substâncias) pensam deveria ser abandonada; e devrígamos recair na expressão comum de que os homens pensam, isto é, de que aquele mesmo ser que, como aparência tem extensão, é (em si mesmo) internamente um sujeito, e não é um composto, mas é simples e pensa.<sup>33</sup>

### **Questões para reflexão**

1. Se a mente e o corpo afetam um ao outro, isso quer dizer que os estados mentais devem realmente ser estados físicos e que o dualismo é falso?
2. A que “eu” se refere na frase “Eu estou com raiva”?
3. Descartes conclui que, se ele está pensando, então ele deve existir como um sujeito incorpóreo, cuja toda essência consiste em pensar. Mas podia ele meramente concluir que há pensamentos ocorrendo?
4. É genuinamente possível conceber a existência de estados mentais independentemente de estados de coisas físicas? Que problemas são colocados pela noção de existência desincorporada?
5. Como é que somos capazes de sentir sensações em nossos membros e corpos e não meramente percebê-los a distância, como um piloto em um navio?

6. Deve haver critérios para distinguir mentes umas das outras, para que a noção de alma possa fazer sentido?

### Sugestões para leituras aprofundadas

- PLATÃO. *Phaedro*, *Meno*, *The Republic*. Há muitas edições disponíveis desses diálogos. Eles contêm os principais argumentos em favor do dualismo, especialmente *Phaedro*. Uma boa exposição e crítica desses argumentos podem ser encontradas em FLEW, A.G.N., *An Introduction to Western Philosophy*. London, Thames and Hudson, 1971.
- DESCARTES, René. *Discourse on the Method*, *Meditations*, *Principles of Philosophy*. Há várias edições disponíveis desses trabalhos.
- Um famoso ataque ao dualismo cartesiano pode ser também encontrado em RYLE, Gilbert. *The Concept of Mind*. Harmondsworth, Penguin, 1973 [1949], que se tornou um trabalho de referência sobre este tópico.
- COTTINGHAM, John. *Descartes*. Oxford, Blackwell, 1986 fornece uma introdução clara e acessível ao dualismo cartesiano, assim como aos principais elementos de sua filosofia. Veja também COTTINGHAM. *A Descartes Dictionary*. Oxford, Blackwell, 1993 para um acesso fácil a uma riqueza de fascinantes detalhes sobre a filosofia de Descartes.
- KENNY, Anthony. *Descartes*. New York, Random House, 1968 é outro livro de utilidade, que contém críticas penetrantes às tentativas de Descartes de estabelecer o dualismo, bem como uma discussão de outros aspectos de sua filosofia. Veja também o excelente estudo de WILSON, Margaret. *Descartes*. London, Routledge, 1976.
- McCULLOCH, Greg. *The Mind and Its World*. London and New York, Routledge, 1995 apresenta uma estimulante introdução ao cartesianismo e aos problemas que ele suscita.
- Entre os filósofos contemporâneos, provavelmente, o caso mais claro e mais detalhadamente argumentado pode ser encontrado em FOSTER, John. *The Immaterial Self*. London and New York, Routledge, 1991. No entanto, o leitor que se inicia em filosofia o achará de difícil compreensão.
- Muito mais breve e mais direto é SWINBURNE, Richard. *Is there a God?* Oxford, Oxford University Press, 1996. c. 5. Veja também seu livro *Evolution of the Soul*. Oxford, Oxford University Press, 1986.
- Uma útil coletânea de extratos dos escritos dos principais filósofos, desde os gregos até o século XX, sobre o problema da mente/corpo é apresentada por FLEW, Anthony, ed. *Body Mind and Death*. Macmillan, Collier Books, 1964.
- SMITH, P. and Jones, O.R. *The Philosophy of Mind*, Cambridge, Cambridge University Press, 1986, além de ser um livro claro e acessível sobre a filosofia da mente em geral, apresenta também uma útil seção sobre o dualismo e seus problemas.

## NOTAS

1. ELIOT, T.S. *Collected Poems [1909-62]* London, Faber & Faber, 1963.
2. ARISTOTLE. "Categories", 2<sup>a</sup> 12. in: SMITH, J.A and ROSS, W.D., eds, *The Works of Aristotle*. Oxford, Clarendon Press, 1910-52.
3. Para uma boa discussão sobre o dualismo de propriedade, veja JACQUETTE, Dale. *Philosophy of Mind*. Engelwood Cliffs, Prentice-Hall, 1994.
4. PLATO. *Phaedo*, 80<sup>a</sup>-b. In: Hamilton, E. and Cairns, H., eds. *The Collected Dialogues of Plato*, Princeton, Princeton University Press, 1961.
5. Os outros argumentos, que o estimulamos a explorar por si próprio, podem ser encontrados principalmente em *Phaedo*, *Meno*, e *Republic*. Uma boa exposição e crítica deles podem ser encontradas em FLEW, A.G.N. *An Introduction to Western Philosophy*. London, Thames and Hudson, 1971, bem como em PRIEST, Stephen. *Theories of the Mind*. Harmondsworth, Penguin, 1991.
6. RYLE, G. *The Concept of Mind*. Harmondsworth, Penguin, 1973 [1949], p. 180-1.
7. A fim de que eu não seja acusado de apresentar uma caricatura rude do cartesianismo, devo assinalar que estou apresentando a interpretação padrão de sua filosofia da mente, que é comumente ensinada no nível "A" e em cursos de graduação. Ela também pode ser encontrada nas reações aos seus trabalhos por parte de filósofos como John Locke (1632-1704) e David Hume (1711-1776) e, mais recentemente, na aclamada obra de Gilbert Ryle *The Concept of Mind*, 1949, à qual devemos a caracterização deliberadamente abusiva do dualismo cartesiano como sendo "o dogma do fantasma na máquina". No entanto, assim como é frequentemente dito que Marx não era ele mesmo marxista, também hoje em dia tem sido afirmado com frequência que Descartes não era um dualista cartesiano! Veja, por exemplo, BAKER, G. and MORRIS, K.J. *Descartes Dualism*. London and New York, Routledge, 1996.
- Há alguma verdade nesta afirmação radical. Por exemplo, mesmo ao tempo em que ele estava quase terminando seu trabalho mais famoso, *Meditations*, Descartes já se estava afastando substancialmente do ponto de vista sobre a natureza e a relação entre mente e corpo que ele havia exposto anteriormente na *Second Meditation*; e maiores refinamentos e diferenças podem ser encontradas em suas cartas e outros trabalhos publicados e não publicados.
- Uma investigação desses outros elementos do pensamento de Descartes não cabe aqui, no entanto. Em conclusão, o dualismo cartesiano como é tradicionalmente compreendido certamente representa uma posição possível – de fato, de acordo com alguns filósofos ela representa o ponto de vista do senso comum do que a mente é – eis aí a razão pela qual me restringi (amplamente) a ele neste capítulo inicial.
8. DESCARTES, René. *Meditations*, London, Everyman Classics, 1986, p. 88.
9. DESCARTES, René. "Carta a um correspondente anônimo. Agosto, 1641". in: GEACH, P and ANSCOMBE, G.E.M. eds. *Descartes: Philosophical Writings*. Nelson University Paperbacks for The Open University, 1970, p. 266.
10. DESCARTES, René. *The Principles of Philosophy*, Secção IX. London, Everyman Classics, 1986, p. 167.
11. Cf. DESCARTES, René. *The Passions of the Soul*. VOSS, S.H. ed. Indianápolis, Hackett, 1989, i, 6, 5.
12. DESCARTES, René. *The Discourse on the Method*. [VEITCH, J., trad.]. London, J.M.Dent & Sons, 1912, p. 44.
13. DESCARTES, René. *The Discourse on the Method*. In: GEACH, P and ANSCOMBE, G.E.M., eds. *Descartes: Philosophical Writings*, p. 41-2.
14. Ibid. p. 41.
15. SPINOZA, B. Ethics. In: CURLEY, E, ed. *The Collected Works of Spinoza*. Princeton, Princeton University Press, 1985, Prefácio à parte III, nota à proposição 2.
16. DESACRTES, René. *Meditations* in: GEACH, P and ANSCOMBE, G.E.M., eds. *Descartes: Philosophical Writings*, p. 117.
17. Ibid.
18. COTTINGHAM, J. *A Descartes Dictionary*. Oxford, Blackwell, 1993, p. 75.
19. DESCARTES, René. *Meditations*. London, Everyman Classics, 1986, p. 86.
20. AYER, A.J. *The Problem of Knowledge*. Harmondsworth, Penguin, 1956, p. 82.
21. DESCARTES, René. *The Discourse on the Method*. In: GEACH, P and ANSCOMBE, G.E.M., eds. *Descartes: Philosophical Writings*, p. 32.

22. HOBES, T. "Second Objection of the Third Set of Objections and Replies containig the controversy between Hobbes and Descartes". in: GEACH, P and ANSCOMBE, G.E.M, eds. *Descartes: Philosophical Writings*, p. 128.
23. Ibid. p. 114-5.
24. KENNY, A. *Descartes*. New York, Random House, 1968, p. 84.
25. WILSON, M.D. *Descartes*. London, Routledge, 1978, p. 193.
26. Apud Ibid.
27. Apud Ibid. p. 196.
28. Apud Ibid. p. 195.
29. COTTINGHAM, J. *Descartes*. London, Phoenix, 1997, p. 35.
30. Apud WILSON. M.D. *Descartes*, p. 179.
31. KANT, I. *Critique of Pure Reason*. Várias edições, A.363.
32. Ibid., A.360.
33. Ibid.

# 3

## A TEORIA DA IDENTIDADE MENTE/CÉREBRO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Fornecer um esboço básico da teoria da identidade mente/cérebro.
- Explicar a diferença entre uma redução ontológica e uma analítica do mental para o físico.
- Explicar as diferenças entre as versões tipo-tipo e espécime-espécime da teoria da identidade.
- Esboçar a evidência que pode dar suporte a uma teoria da identidade e compreender os pontos fortes da teoria.
- Explicar e acessar várias objeções à teoria da identidade.

### Uma breve retrospectiva histórica

Como vimos no último capítulo, o dualismo afirma que os eventos mentais nunca podem ser idênticos aos eventos físicos. Os eventos mentais são completamente não-físicos em todos seus aspectos e consistem em mudanças nos estados não-físicos de uma entidade imaterial, a alma.

A teoria da identidade mente/cérebro nega o que o dualismo afirma. A mente não é uma entidade separada e distinta do cérebro. Fazendo uma afirmação direta e simples, a mente é idêntica ao cérebro

vivo, e eventos mentais são simplesmente eventos cerebrais. Não há substâncias fantasmáticas imateriais ou eventos para constituir nossas mentes. Tudo, incluindo nossa própria consciência e pensamentos, é puramente material ou físico. A teoria, portanto, exemplifica o materialismo.

A teoria da identidade mente/cérebro tornou-se proeminente no início da década de 1950 e na seguinte, e foi advogada por filósofos tais como U. T. Place, J. J. C. Smart e D. Armstrong. No entanto, o materialismo tem uma longa história. O filósofo da Grécia antiga, Epicuro (342-270 a.C.), graficamente expressou o con-

ceito de materialismo quando disse: "Por convenção a cor existe, por convenção o amargo, por convenção o doce, mas, na realidade, átomos e vazio". Tudo que em última instância existe são átomos materiais rodopiando no espaço. Estes se juntam, temporariamente, para constituir coisas materiais como a terra e seus acessórios, e organismos físicos, incluindo animais humanos e não-humanos. No entanto, na vastidão da eternidade, tudo é evanescente, fugidio, e tudo retornará ao vazio para re-formar, a partir de éons lentos, novas configurações de partículas físicas, que, por sua vez, eventualmente se dissolverão. Para os dualistas, a vida após a morte é uma possibilidade genuína porque, como substância lógica, a alma pode existir por si mesma, independentemente do corpo. Entretanto, para o materialista epicurista, a mente é inteiramente uma parte do mundo material, e termina quando o resto do corpo termina. Não há, porém, necessidade alguma de temer a morte, porque, como um seguidor de Epicuro, o pensador romano Lucrécio (c.98-c.55 a.C.) faz a picante colocação: "Quando estamos lá, a morte não está; quando ela está lá, nós não estamos".

Passando adiante, nos últimos 400 anos numerosos pensadores reagiram contra o dualismo cartesiano e abraçaram o materialismo de diferentes complexões. Thomas Hobbes (1588-1679) mantinha que a noção de uma substância incorpórea é uma contradição em termos, e também insistia que "conceitos e aparições não são realmente nada mais do que movimentos em alguma substância interna à cabeça". John Locke (1632-1704) fez a concessão, naquele tempo ousada e perigosa, de que Deus, se assim quisesse, poderia adicionar

pensamento aos movimentos da matéria. O ateu Barão d'Holbach, em *The System of Nature*\*, escreveu sobre o erro que cometeu o homem quando:

Ele fez a conjectura de que ele não era somente um ser distinto, mas de que ele estava separado, com energias diferentes, dos outros seres na natureza; que ele era de uma essência mais simples, [isto é, a alma cartesian] não tendo nada em comum com qualquer coisa pela qual fosse cercado; nada que o conectasse com tudo aquilo que ele visse.

Ainda outro materialista foi Julien Offray de La Mettrie (1709-1751), como os títulos de seus dois livros *Man a Machine*\*\* e *Man a Plant*\*\*\* proclamam claramente. La Mettrie afirmava que o homem é um produto evolutivo da natureza assim como os demais animais – não há ponto algum de separação nítida entre nós e os animais não-humanos, contrariamente ao que Descartes asseverava. Os seres humanos são essencialmente máquinas controladas pela neurologia do cérebro: "O pensamento está tão longe de ser incompatível com a matéria organizada que a mim parece ser somente uma outra de suas propriedades, tais como a eletricidade, a faculdade motora, a impenetrabilidade, a extensão, etc.".

A semelhança de algumas das formas de materialismo anteriormente aqui esboçadas, a teoria da identidade mente/cérebro adota uma atitude descompromissada no que diz respeito à crença em mentes concebidas como de natureza não-física. Os fenômenos mentais não dispõem de uma existência sobre e acima dos fenômenos físicos: eles são fenômenos físicos. A teoria da identidade exemplifica o monis-

\*N. de T.: Em português: *O sistema da natureza*.

\*\* N. de T.: Em português: *Homem uma máquina*.

\*\*\* N. de T.: Em português: *Homem uma planta*.

mo, o monismo material, para ser preciso, porque diz que somente substâncias materiais e seus estados existem. Pode também ser dito que exemplificam o fisicalismo.

O fisicalismo diz que os seres humanos são entidades inteiramente materiais, cujo funcionamento e propriedades podem ser completamente explicadas pelos conceitos e teorias extraídas de uma física idealmente completa. Não há espaço para invenções imateriais ou sobrenaturais nas cadeias causais físicas que correm pelo sistema nervoso central de uma pessoa, que vão desde *inputs* sob a forma de estímulos físicos até a produção de comportamentos. Os seres humanos são inteiramente partes do mundo físico natural, não fantasmas sobrenaturais inexplicavelmente ligados a máquinas corpóreas. O que quer que seja a mente, ela não pode ser imaterial, e alguma expli-cação que a coloque inteiramente dentro do mundo natural deve ser possível.

### O que a teoria da identidade afirma e o que não afirma

É importante ser claro quanto ao que precisamente os teóricos da identidade estavam afirmando, e o que não estavam, quando disseram que os estados mentais eram estados cerebrais. Para começar, eles não estavam dizendo que falar sobre estados mentais tinha o mesmo significado de falar sobre estados cerebrais. Suponha, para tomar emprestado um exemplo muito utilizado na literatura, que dores são teoricamente ditas serem idênticas a fibras C disparando\* no sistema nervoso. Bem, os teóricos da identidade nunca desejaram dizer que a frase "Estou com dor" fosse sinônima de "Minhas fibras C estão disparando". Dizer que "Estou com dor, mas minhas fibras C não estão disparando" pode ser falso, mas não é uma contra-

dição em termos, como seria caso ambas as metades da frase fossem equivalentes quanto ao seu significado. A afirmação de que "Dores são idênticas a fibras C disparando" **não** é semelhante a "Todos os triângulos são idênticos a figuras de três lados". Isso porque "trilateral" tem o mesmo significado de "figura de três lados", então descrever algo como um trilateral, mas negar que tenha três lados seria formalmente contradizer-se. "Trilaterais são figuras com três lados" é analiticamente verdadeira, isto é, necessariamente verdade em virtude dos significados dos termos que contém, e nunca pode ser falsa. Portanto, toda conversa sobre trilaterais pode ser colocada, sem qualquer prejuízo de significado, em conversa sobre figuras de três lados. Isso é o que os filósofos chamam de redução analítica.

### Evitando objeções iniciais à teoria da identidade

A fim de descartar desde o início algumas objeções *prima facie* à posição deles, os teóricos da identidade foram muito cuidadosos em salientar que a identidade que estavam defendendo não era analítica.

Uma objeção pode serposta da seguinte forma: "O significado de falar sobre meus estados mentais é inteiramente diferente de falar sobre meus estados cerebrais; consequentemente, meus estados mentais não podem ser possivelmente idênticos aos meus estados cerebrais". No entanto, responde o teórico da identidade, a partir do fato de que o significado de falar sobre estados mentais não é o mesmo que o significado de falar sobre estados cerebrais, você não pode deduzir que estados mentais não possam ser estados cerebrais. Considere um exemplo paralelo: O significado da expressão "estrela

\*N. de T.: Em inglês: *firing*, que pode também ser traduzido por "estimulando".

"matutina" não é mesmo significado da expressão "estrela vespertina". Bastante obviamente, a primeira expressão significa a estrela que aparece pela manhã, e a segunda, a estrela que aparece ao anoitecer. Mas isso não pode significar que a estrela matutina não seja idêntica à estrela vespertina. De fato, o planeta chamado por essas expressões é um, e o mesmo, o tempo todo, qual seja, o planeta Vênus. A não sinônima de expressões flanqueando um sinal de identidade não descarta automaticamente a verdade da afirmação de identidade. "O repórter que trabalha com Lois Lane no Daily Planet" não significa a mesma coisa que "O homem de aço". No entanto, o repórter e o homem de aço são idênticos, uma e a mesma pessoa, qual seja, o bom velho Clark Kent.

Um argumento que se relaciona a este diz: "Eu tenho sido capaz de falar sobre meus estados mentais por muitos anos sem saber coisa alguma sobre estados cerebrais, ou mesmo saber se tenho um cérebro ou não. Portanto, meus estados mentais não podem ser idênticos aos meus estados cerebrais". A fraqueza deste argumento se apresenta rapidamente. Eu posso falar de água sem saber coisa alguma a respeito de sua constituição interna, mas isso não pode mostrar que não seja idêntica a coleções de moléculas de  $H_2O$ . Uma nação de dorminhocos, para tomar emprestado um exemplo de J. J. C. Smart,<sup>1</sup> poderia

conhecer tudo sobre a estrela vespertina e absolutamente nada sobre a estrela matutina, uma vez que estaria ferrada no sono durante o alvorecer, mas isso não poderia possivelmente mostrar que a estrela vespertina não fosse uma e a mesma entidade que a estrela matutina.

Isso tudo se torna mais claro quando o significado de uma expressão é diferenciado de sua referência, o que ele nomeia ou refere. No caso de Vênus, esta distinção pode ser claramente ilustrada com um diagrama, como mostra a Figura 3.1. Claramente, seria possível falar sobre a estrela matutina e enumerar qualquer quantidade de fatos a seu respeito, e também falar sobre a estrela vespertina e mencionar qualquer quantidade de fatos a seu respeito, embora não lograsse reconhecer que ambos os domínios de discurso tinham um mesmo referente, e que, de fato, estamos falando sobre um e o mesmo objeto. Isso é o que se aplica no caso da mente e do cérebro, de acordo com os teóricos da identidade. Ao falar sobre eventos cerebrais físicos por um lado e eventos mentais por outro, estamos falando, sem que saibamos, sobre um e o mesmo conjunto de eventos, descritíveis em dois vocabulários distintos. Assim sendo, parece que conhecemos conjuntos diferentes de fatos sobre dois conjuntos de eventos diferentes, enquanto a verdade é que os fatos estão relacionados a uma úni-

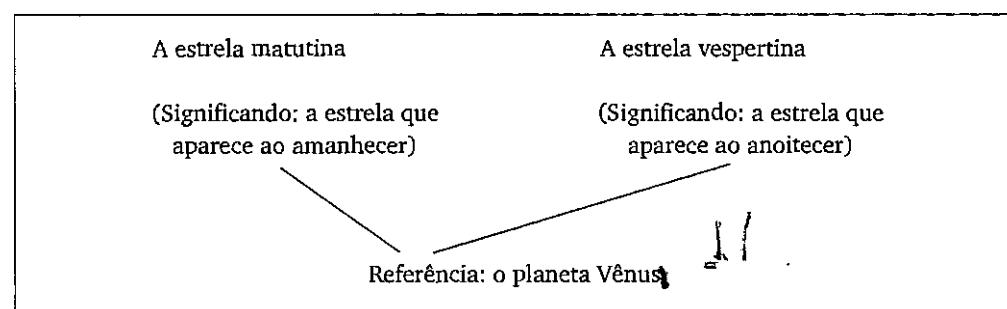


Figura 3.1 Significado e referência.

ca realidade descritível tanto em termos mentais quanto físicos. O erro básico que sustenta este argumento contra a teoria da identidade é a tentativa de descobrir a verdade sobre a realidade com base naquilo que se sabe ou não se sabe. No entanto, esta é realmente uma base muito pobre para se alcançar conclusões substantivas sobre a natureza da mente. Caso os estados mentais fossem idênticos aos estados cerebrais, o simples fato de não se reconhecer que sempre que fatos mentais fossem referidos a alguém estivessem de fato referindo também estados cerebrais não poderia possivelmente mostrar que a identidade mente/cérebro não foi alcançada.

### A teoria da identidade tipo-tipo

Somente a identidade de coisas individuais está envolvida em afirmações tais como a estrela matutina é a estrela vespertina, ou Clark Kent é o Super-Homem, e isto não foi de especial interesse para os teóricos da identidade, simplesmente fornecendo alguns exemplos simples que poderiam ser usados para esboçar a simples ossatura da teoria que estavam propondo sobre a natureza da mente.

Inspirando-se na ciência, seu objetivo real era constituir um caso para a identidade de classes de fenômenos – quais sejam, as de fenômenos mentais em geral – com determinadas formas de processos físicos no cérebro. Os cientistas haviam obtido alguns êxitos espetaculares na

identificação de uma classe de fenômenos físicos com outra classe, e havia esperança de que o mesmo pudesse ser alcançado na identificação dos fenômenos mentais com os físicos.

Para que o devido valor seja atribuído a esses avanços, tem que ser lembrado que, antes do advento da ciência moderna, a natureza de diversos fenômenos tais como os relâmpagos, os genes e a água não era compreendida. Por exemplo, houve época em que se acreditava que os relâmpagos fossem raios lançados pelos deuses. Agora sabemos que os relâmpagos constituem um padrão de descargas elétricas. O primeiro filósofo ocidental conhecido, Thales, da Grécia antiga (c.600 a.C.), acreditava que a água fosse um elemento básico, a partir do qual todas as demais coisas fossem feitas. No entanto, os avanços científicos demonstraram que a água não é um elemento, mas um composto formado de hidrogênio e oxigênio. Havia a hipótese de que os genes seriam tudo aquilo que fosse responsável pela transmissão de traços genéticos, mas até que Crick e Watson fizessem seu trabalho desbravador no século XX, ninguém sabia que os genes consistem em sequências codificadas de moléculas de DNA. Resumindo, podemos ver que esses exemplos ilustram identidades científicas, que podem ser expressas de acordo com a Figura 3.2. Os termos do lado direito da equação devem ser pensados como reveladores da natureza interna, escondida dos itens constantes à esquerda. Além disso, estas identidades são universalmente obtidas: elas não se aplicam meramente a indi-

Água =  $H_2O$

Relâmpago = um padrão de descarga elétrica

Genes = sequências codificadas em moléculas de DNA

Figura 3.2 Identidades científicas.

víduos, instâncias isoladas, mas a tipos. O tipo de coisa água é idêntico ao tipo de coisa  $H_2O$ . O clarão do relâmpago que agora experiencio não será o único clarão de relâmpago que constitui um padrão de descargas elétricas. Todos os clarões de relâmpagos, em todos os lugares, em qualquer tempo, também constituirão padrões de descargas elétricas. O tipo de fenômeno conhecido como clarão de relâmpago é idêntico ao tipo de evento que consiste em um padrão de descarga elétrica.

Assim como as identidades científicas listadas, a teoria da identidade tipo-tipo afirma que cada tipo de estado mental será idêntico a um dado tipo de estado cerebral. Assim, por exemplo, todas as vezes que alguém, em qualquer lugar, tiver uma dor, seu cérebro estará em um tipo particular de estado físico, do mesmo modo que, a qualquer hora e lugar que alguém encontra o tipo de coisa água, estará também encontrando o tipo de coisa  $H_2O$ . Qual tipo de estado cerebral virá a ser idêntico a estar com dor não pode ser especificado antecipadamente, mas terá de ser descoberto empiricamente pela observação dos estados cerebrais das pessoas que estão com dor. Se uma correlação de um a um, entre estar com dor e existir algum tipo dado de estado cerebral, é encontrada em todos os casos observados, o teórico da identidade mente/cérebro argumentará que a melhor explicação da correlação é a identidade dos itens correlacionados. Uma

representação diagramática da identidade tipo-tipo pode ser vista na Figura 3.3.

### A teoria da identidade tipo-tipo e o reducionismo

Como testemunhamos, a teoria tipo-tipo rejeita a redução analítica do mental ao físico. Mas ela argumenta que é razoável crer que eventos que caiam em um determinado tipo de descrição mental devam sempre cair no mesmo tipo de descrição física. Uma outra forma de colocar isto é dizer que os significados de termos mentais e físicos são diferentes, mas que suas extensões – os estados que entram sob conceitos mentais e físicos – coincidem exatamente. Dizendo mais precisamente, a teoria afirma que, se um estado mental M é idêntico a um estado mental B, M pode manter-se somente se B se mantém, e B pode manter-se somente se M se mantém. Isso é exatamente o que devemos esperar caso M seja idêntico a B. Para dar um exemplo incontroverso, a água pode ser encontrada em um dado lugar se, e somente se,  $H_2O$  pode ser encontrado no mesmo lugar. Isso é o que os filósofos significam por redução ontológica – um grupo de fenômenos que é aparentemente numericamente diferente de outro grupo de fenômenos vem a ser simplesmente *um* conjunto de existentes, e não dois. A expressão “se, e somente se” é conhecida como bicondicional (porque

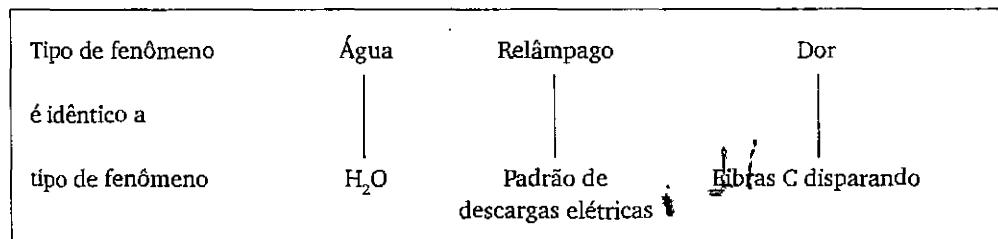


Figura 3.3 Identidade tipo-tipo.



ela contém dois “ses”). A esperança do teórico tipo-tipo é a de que as leis da ponte bicondicional – leis psicofísicas – que fazem a conexão de tipos de estados mentais com determinados tipos de estados físicos estejam a caminho. Um exemplo de uma dessas leis, que usa a identificação de dores com fibras C disparando, teria a seguinte leitura: a dor existirá se, e somente se, as fibras C estiverem disparando em um sistema nervoso central humano. Em outras palavras, o disparo de fibras C é logicamente necessário e suficiente para que M ocorra.

A que isso começa a conduzir é a afirmação de que os eventos mentais não só são idênticos aos eventos físicos, mas que as propriedades mentais, em virtude das quais os eventos são mentais, se mostrariam idênticas às propriedades físicas em virtude das quais os eventos físicos são físicos. Os conceitos mentais permanecem não-sinônimos de conceitos físicos, mas as classes de propriedades aparentemente diferentes, mental e física, que entram em cada tipo de conceito, respectivamente, se mostrariam como nada mais do que uma única propriedade, que pode ser descrita por meio de dois vocabulários distintos.

Para tornar isso mais claro, compare a situação com o que diz respeito às propriedades de um gás. Falar a respeito da temperatura de um gás não é sinônimo de falar sobre a energia cinética média de suas moléculas. Negar que a temperatura de um gás seja idêntica à energia cinética média de suas moléculas pode ser falso, mas não é autocontraditório. No entanto, a temperatura do gás não é nada além da energia cinética média de suas moléculas. Não há dois conjuntos de fatos separados e independentes aqui – fatos de temperatura e fatos de energia cinética média de moléculas –, mas um estado de coisas descritível de dois modos diferentes. De fato,

o que aconteceu no caso do gás é uma redução mais ou menos suave do macrofenômeno da temperatura ao microfenômeno envolvido no comportamento das moléculas. Para que tal redução seja possível, deve haver, como um simples mínimo, um conjunto de correlações, sem exceção entre os valores específicos para a temperatura do gás e os valores específicos para a energia cinética molecular média. A menos que haja uma relação sistemática entre esses dois conjuntos de variáveis, de modo a nos fornecer uma boa razão para acreditar que um dado valor de uma variável possa se manter se, e somente se, um dado valor se mantém bom para a outra variável, a perspectiva de uma redução de um fenômeno ao outro é descartada, porque a ausência de uma correlação invariável abre a possibilidade de que um fenômeno possa variar independentemente do outro. O objetivo da redução é demonstrar que a verdadeira natureza do fenômeno reduzido é fornecida pela base à qual ele está reduzido. A verdadeira natureza da temperatura de um gás, supostamente, consiste em fatos sobre a energia cinética média de suas moléculas, mas isso não poderia ser o caso se a temperatura variasse independentemente da energia cinética das moléculas do gás.

O teórico da identidade tipo-tipo argumenta que as propriedades mentais são passíveis de uma redução de modo semelhante, uma vez que um conjunto de correlações sem exceção entre um dado tipo de estado mental e um dado tipo de estado cerebral esteja a caminho. Será, então, razoável concluir que a dolorosa sensação de uma dor, por exemplo, é idêntica ao comportamento de determinados neurônios no sistema nervoso central, muito embora falar sobre dores não tem o significado equivalente ao de falar sobre o comportamento de neurônios.

O tipo de observação que dá esperança à teoria tipo-tipo é instanciado pelas descobertas de escâneres PET\* – tomografia de emissão de pósitrons – e MRI\*\* – ressonância magnética funcional por imagem. O que a pesquisa tem revelado é que o exercício de certas capacidades mentais envolve com certeza regiões distintas do cérebro, que “acendem-se” quando funções específicas são executadas. Por exemplo, as áreas do cérebro que mostram um acréscimo de atividade quando uma pessoa fala e ouve palavras concordam exatamente com as áreas da linguagem identificadas por Broca e Wernicke ao menos 100 anos antes que surgissem as modernas técnicas de escaneamentos. Se todo o cérebro pudesse ser mapeado de modo semelhante a esse, mostrando correspondências de um a um entre certos tipos de atividade mental e neurológica, o caso para a redução tipo-tipo da atividade mental para a fisiológica se tornaria mais forte.

### A teoria da identidade espécime-espécime e a tese da múltipla capacidade de realização\*\*\*

A teoria tipo-tipo insiste que cada tipo de estado mental resultará idêntico a um tipo dado de estado físico, e esses não serão exceção à regra. Isso faz com que essa hipótese seja muito forte e exigente. No entanto, enquanto permitimos que a água seja sempre  $H_2O$ , somos igualmente levados a aceitar que a dor, por exemplo, onde quer que ela ocorra e a qualquer hora, deva ser identificada com fibras C disparando? Não poderia a dor, em algumas ocasiões, em pessoas diferentes, ou

em animais, ser idêntica a um tipo distinto de estado físico? É essa ideia, a ideia de que estados mentais poderiam se dar de formas múltiplas – isto é, incorporado em todos os tipos de arranjos materiais, assim como um programa de computador pode ser realizado em uma variedade de meios, tais como discos rígidos ou flexíveis, chips de silício e assim por diante – que levaram à demissão da teoria tipo-tipo no início dos anos 1970, não muito depois do seu surgimento.

A teoria espécime-espécime tomou essa ideia e a desenvolveu em uma crítica importante da teoria tipo-tipo. Para que se compreenda como ela consegue fazer isso, no entanto, devo em primeiro lugar explicar a diferença entre um tipo e um espécime. A distinção pode ser facilmente ilustrada em referência à expressão “tipo-tipo”. Esta expressão contém quatro tipos de letras, quais sejam: “t”, “i”, “p” e “o”. Mas cada uma dessas letras ocorre duas vezes na expressão: isto é, há duas instâncias, ou espécimes, de “t”, duas instâncias de “i” e assim por diante. Claramente, se a água, considerada como um tipo, é idêntica ao tipo de coisa  $H_2O$ , então cada instância individual ou espécime de água, uma poça de água da chuva no meu jardim dos fundos, por exemplo, deve ser idêntica a uma instância ou espécime de  $H_2O$ . A teoria da identidade tipo-tipo então implica a identidade dos espécimes dos tipos em questão. No entanto, o contrário não se mantém. Cada espécime de um tipo de estado mental poderia ser idêntico a um espécime de um tipo de estado físico, mas não necessita ser o caso que espécimes do mesmo tipo de estado físico devam ser envolvidos a cada ocasião. Compare o que é verdadeiro no caso de relógios. Cada

\*N. de T.: Por extenso em inglês: *Positron Emission Tomography*.

\*\*N. de T.: Por extenso em inglês: *Functional Magnetic Resonance*.

\*\*\*N. de T.: No original: *realizability*.

relógio será idêntico a algum arranjo de partes, mas claramente seria uma afirmação ridícula a de que cada espécime do tipo "relógio" deva ser idêntico a um espécime do mesmo tipo de arranjo de partes exatamente. Os relógios são feitos em uma variedade de formas, tamanhos e mecanismos, embora, ao final de contas, eles sejam todos idênticos a arranjos de materiais.

Do mesmo modo, cada espécime do pensamento-tipo "Eu coloquei o gato para fora?" será idêntico a algum espécime de estado cerebral ou outro, mas não segue-se que será sempre idêntico a um espécime do mesmo tipo de estado cerebral. Talvez o espécime de pensamento sobre o gato *em mim* seja idêntico a um espécime do tipo de processo cerebral no qual fibras C disparam, mas *em você* o mesmo espécime de pensamento seja idêntico a um espécime do tipo de processo cerebral no qual fibras Z disparam. De fato, mesmo *em mim* é concebível que espécimes

do mesmo tipo de pensamento possam ser idênticos a espécimes de diferentes tipos de processos cerebrais em ocasiões distintas. O que acontece no caso de vítimas de ataque cardiovascular dá suporte a esse ponto. Uma parte do cérebro é danificada e a vítima perde a capacidade de caminhar. Mas, então, uma outra parte do cérebro gradualmente assume aquela função, e a capacidade de caminhar é restaurada. O mesmo pode ser aplicado aos pensamentos. Sempre que eu pensasse sobre colocar o gato para fora, antes do meu ataque cardiovascular, as fibras C disparariam. No entanto, minhas fibras C estavam irreparavelmente danificadas, e agora minhas fibras Z disparam no lugar daquelas quando eu tenho o pensamento sobre o gato, estas fibras havendo tomado a função das fibras C originais. A Figura 3.4 ilustra a teoria espécime-espécime (as letras minúsculas representam espécimes, as maiúsculas, tipos).

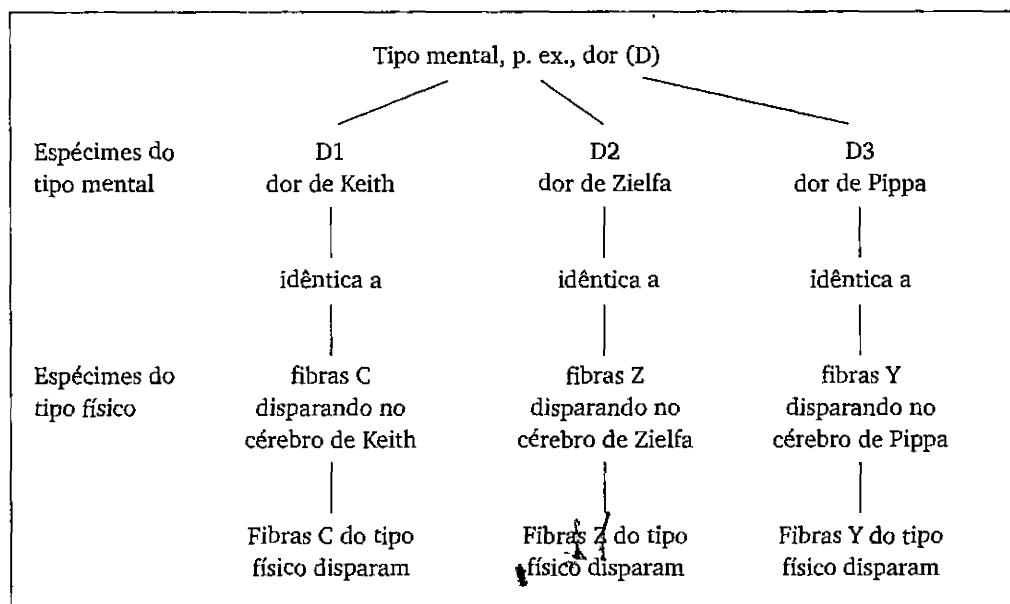


Figura 3.4 A teoria espécime-espécime.

Qual teoria, a tipo-tipo ou a espécime-espécime, tem mais probabilidade de ser verdadeira? (Certamente, ambas podem ser falsas, se o dualismo for verdadeiro, mas não estamos considerando essa possibilidade por agora.)

A plasticidade do cérebro, o fato de que diferentes partes do cérebro podem assumir as funções que outras partes abandonaram, contam fortemente contra a teoria tipo-tipo. Além disso, a teoria tipo-tipo parece ser desnecessariamente restritiva. Por que a mentalidade deve ser confinada a cérebros humanos e à biologia humana? Por que ela não pode ser incorporada em sistemas físicos completamente diferentes, como a biologia de baleias, ou alienígenas, ou até mesmo sistemas inorgânicos encontrados em robôs e andróides? Somente a experiência, e não a teorização filosófica antecipada, poderá nos dizer se essas possibilidades são genuínas.

Insistir que a mente só pode se manifestar em cérebros humanos tem sido etiquetada como uma manifestação de chauvinismo, um comprometimento infundado e fanático à crença de que a mentalidade não pode ser encontrada em nenhum lugar senão na neurologia humana. Voltaremos mais tarde a essa questão, no Capítulo 5, quando trataremos do funcionalismo como teoria da mente.

Em resposta à teoria espécime-espécime foi argumentado que podemos manter a teoria tipo-tipo, uma vez que seu escopo seja restringido, a fim de que uma versão particular dela seja concebida para manter-se somente dentro de uma dada espécie. Assim, a dor em seres humanos sempre consistirá em neurônios P disparando, mas em símios ela sempre será neurônios X, em gatos será invariavelmente neurônios Y, e assim por diante, em todas as diferentes espécies de criaturas ou seres dotados de mentes.

No entanto, não há garantia alguma de que mesmo uma forma restrita da identidade de tipo será verdadeira. Como salientei anteriormente, por que não poderia ser verdade que quando você e eu pensamos sobre o gato, tipos totalmente diferentes de neurônios disparem no meu caso, quando comparados aos seus? Indo mais adiante, mesmo no caso de um único indivíduo, por que, de um momento a outro ao longo da vida daquela pessoa, deve ser o mesmo tipo de neurônios que é ativado quando aquela pessoa tem certo tipo de pensamento? A teoria espécime-espécime, consequentemente, parece ser mais plausível do que a teoria tipo-tipo. No entanto, como veremos quando examinarmos as objeções à teoria da identidade, a adoção da teoria espécime-espécime ameaça minar todo o empreendimento da identificação do mental com o físico.

### Vantagens da teoria da identidade

A teoria da identidade, caso alguma de suas versões possa ser posta em marcha, é uma doutrina atraente por uma série de questões. Em primeiro lugar, é uma teoria comparativamente simples – os estados mentais são simplesmente estados cerebrais. Tudo mais sendo igual, a teoria mais simples, que faça o menor número de pressupostos e possa ser formulada postulando o menor número de entidades ou propriedades, é a teoria preferida. Diferentemente do dualismo, que tem de afirmar tanto substâncias materiais quanto imateriais, e o materialismo de propriedades, que tem de postular tanto propriedades físicas quanto não-físicas, a teoria da identidade pode sair-se com somente coisas físicas e propriedades físicas.

Em segundo lugar, o problema de Descartes referente a como a mente pode

afetar o corpo simplesmente desaparece. Se os estados mentais são simplesmente estados cerebrais, então o problema da interação mente/corpo é eliminado, porque o que inicialmente parece ser mental para a causação física vem a ser, ao final de contas, meramente físico para a causação física. Os eventos físicos no cérebro formam parte de uma cadeia física que se estende a partir de estímulos físicos do meio ambiente da pessoa, que impingem sobre os sentidos e causam mudanças no cérebro da pessoa, e esses, por sua vez, produzindo uma resposta comportamental, que causa mudanças no mundo além da pele da pessoa. Enquanto não entendermos inteiramente a física como causalidade física, ao menos estamos eximidos da tarefa de explicar como uma substância não-física, desprovida de dimensões, teria a possibilidade de ter um efeito sobre a matéria extensa.

Em terceiro lugar, não somos capazes de compreender instantaneamente por que mudanças no cérebro devidas a danos, doença, operações e drogas são acompanhadas de alterações no funcionamento mental. Ao alterar o cérebro, estamos alterando a mente, porque a mente é justamente o cérebro. De modo semelhante, podemos também verificar por que o aumento no tamanho do cérebro e sua complexidade, à medida que subimos na escala evolucionária, estão emparelhados em todo o caminho ascendente a aumentos correspondentes na sofisticação da capacidade intelectual e na riqueza da vida mental.

### **Problemas para a teoria da identidade**

Infelizmente, ainda não estamos em casa e sequinhos ainda. Há ainda uma quantidade de dificuldades sérias às quais a teoria da identidade tem de superar, em

qualquer uma de suas versões, para que possa ser aceita como a solução para o enigma mente/corpo. Tais dificuldades são longamente discutidas nesta seção.

#### **O mental e o espacial**

Esta objeção à teoria da identidade mente/corpo, que veremos em cumprimento a uma promessa feita no Capítulo 1, seção “Há um traço do mental?”, inspira-se essencialmente em Descartes, que dividiu o mundo em duas categorias mutuamente exclusivas: o domínio da *res extensa* – ampliado, coisas que ocupam espaço – e o mundo da *res cogitans* – o domínio das coisas conscientes não-extensas. A localização espacial dos processos cerebrais pode ser especificada com alto grau de precisão. Escâneres podem mostrar com riqueza de detalhes aquelas áreas do cérebro que se tornam ativas quando os sujeitos são solicitados a realizar várias tarefas intelectuais e práticas. No entanto, parece não fazer sentido algum especificar a localização de um pensamento ou de uma emoção. Afirmar que o pensamento “Coloquei o gato para fora?” está localizado bem atrás de seus olhos, ou que uma pontada de ciúme pode ser encontrada na região occipital do cérebro, na parte traseira da cabeça, é dizer coisas sem sentido. Além disso, enquanto falamos de processos cerebrais como rápidos ou lentos, ou movimentos moleculares no cérebro como retos ou circulares, não faz sentido algum dizer isso de minhas crenças ou sentimentos. Portanto, estados mentais não podem ser identificados com processos cerebrais.

Este argumento contra a teoria da identidade parece obter sucesso à primeira vista. Na realidade, no entanto, ela incorpora certas confusões, que, uma vez esclarecidas, o tornam muito menos conclusivo. Para que se compreenda o que há

de errado com ele, necessitamos voltar à noção de uma substância lógica, que foi caracterizada no início do Capítulo 2, (seção “Dualismo substancial”). Substâncias lógicas são entidades que podem existir por si mesmas, independentemente de qualquer outra coisa, e certamente faz sentido perguntar em que lugar no espaço essas entidades estão localizadas. O cérebro, como uma substância lógica composta, e as várias partes do cérebro, incluindo os microelementos que o compõem, têm uma locação no espaço mais ou menos confiável. Os processos cerebrais, que consistem em movimentos de substâncias lógicas tais como moléculas, também ocupam uma região no espaço: o movimento de um grupo de moléculas complexas de uma região do cérebro para outra poderia ser rastreada de uma área para outra, e sua trajetória espacial poderia ser delineada. No entanto, segundo a queixa contra a teoria da identidade, pensamentos, sensações, emoções, crenças e outros estados mentais não podem uma literal localização no espaço atribuído a eles, diferentemente de regiões e componentes do cérebro. Consequentemente, não há meio de o mental e o físico virem a ser idênticos.

A objeção, no entanto, se baseia em um sério engano do estatuto ontológico de vários itens mentais citados. Por meio da caracterização de tais itens usando expressões semelhantes a substantivos, é fatalmente fácil pensar estados mentais como substâncias lógicas. Concebidos dessa forma, parece que esses devem ter uma localização específica no espaço, juntamente com outras substâncias lógicas; no entanto, quando tentamos dizer qual é a localização neste caso de estados mentais, acabamos dizendo coisas sem sentido. Por sua vez, isso leva à negação de que estados mentais tenham qualquer localização espacial literal, e, consequentemente, sua identificação com estados mentais, que

efetivamente possuem uma localização espacial, é proibida desde o início.

Para ver o que há de errado nessa linha de raciocínio, no entanto, temos de lembrar que dores, crenças, pensamentos e emoções (e estados mentais em geral) não são substâncias lógicas. Embora as pessoas tenham dores, pensamentos, crenças e emoções, isso não denota uma relação entre as duas entidades – uma pessoa, por um lado, e uma crença ou uma dor, por outro –, mas sim uma condição de uma única coisa. Temos uma pessoa que está sentindo uma dor, ou pensando sobre alguma coisa ou acreditando que algo seja verdadeiro. Agora, essas coisas, necessariamente, continuarão independentemente de quem sejam as pessoas. Kekulé estava em cima de um ônibus quando ele pensou sobre a estrutura da molécula de benzeno; assim, aquele foi o lugar onde seus pensamentos sobre o benzeno ocorreram. Semelhantemente, eu estava na minha cama, nas primeiras horas da manhã, quando fui acometido por severas dores que anunciam envenenamento sanguíneo, alguns anos atrás; assim, lá foi o lugar onde minhas dores ocorreram. Pedir uma localização mais precisa para os pensamentos de Kekulé do que o topo do ônibus no qual estava sentado, pedir onde precisamente dentro dele os pensamentos estavam ocorrendo, já é fazer uma demanda ilegítima, que resulta em nada menos do que querer atribuir a pensamentos o status de substâncias lógicas. O mesmo poderia ser verdadeiro, mesmo no caso de minhas dores, porque embora eu pudesse referir intensas dores no meu lado, não segue que há coisas chamadas dores que existam dentro do meu corpo, como meus nervos e músculos o fazem. De fato, há somente eu, uma entidade individual, na condição de estar com dor.

A que isso está conduzindo é a um aperfeiçoamento da teoria da identidade,

que pode contornar a objeção originalmente feita a ela. O primeiro a fazer essa formulação foi o filósofo americano Thomas Nagel em um escrito chamado "Physicalism"<sup>\*</sup>. Nagel argumenta que, se os dois lados da identidade proposta fossem ajustados de tal forma que se mantivessem, não entre um processo cerebral e a ocorrência de uma dor ou um pensamento, mas sim entre "o meu ter uma certa sensação e o *estar* de meu corpo em um certo estado físico, então ambos estarão indo ao mesmo lugar – qual seja, onde quer que eu (e meu corpo) ocasionalmente estejamos [...] mesmo se uma dor esteja localizada na minha bochecha direita, eu estou tendo aquela dor no meu escritório, na universidade".<sup>2</sup> Uma vez feito o ajuste de Nagel, ficará aparente que o argumento a partir da alegada não-espacialidade do mental contra a teoria da identidade perde muito de sua força, senão toda ela.

### **A simetria de afirmações de identidade**

A segunda dificuldade para a teoria da identidade que examinarei éposta pela simetria de afirmações de identidade. Se A é idêntico a B, então B é idêntico a A. Isso quer dizer, tomando um exemplo anterior, que se a estrela matutina é idêntica à estrela vespertina, então, igualmente, a estrela vespertina deve ser a estrela matutina. Isso parecerá não somente óbvio, mas talvez terrivelmente trivial. No entanto, a importante lição a ser aprendida é a de que, devido à simetria da identidade, não podemos argumentar que Clark Kent seja realmente o Super-Homem (realmente, verdadeiramente), mas que o Super-Homem não é realmente Clark Kent. Ou, tomando um exemplo de uma identidade científica, que a água seja realmente H<sub>2</sub>O, mas que H<sub>2</sub>O não seja realmente água.<sup>3</sup>

Antes de prosseguir, veja o Exercício 3.1.

### **EXERCÍCIO 3.1**

Tente aplicar a simetria da identidade à teoria da identidade mente/corpo. Que resultado você obtém? Qual significância você pensa que isso tem?

O resultado é que você não pode dizer que estados mentais sejam realmente só processos cerebrais (realmente, verdadeiramente), mas que processos cerebrais não são realmente estados mentais. Fazer esta afirmação privilegia o físico sobre o mental sem justificativa. O que ela faz, de fato, é livrar-se de traços mentais em favor de traços cerebrais. Ela afirma, efetivamente, que, ao final de contas, só existem traços cerebrais. Propriedades mentais e estados foram neutralizados, isto é, o mental foi assimilado e reduzido ao puramente físico. Certamente, len-do a afirmação de identidade da direita para a esquerda conduz a uma absurdade igual e oposta, qual seja, o erro de mentalizar o cérebro. Segundo essa leitura, os processos cerebrais são realmente somente processos mentais (realmente, verdadeiramente) e o físico foi absorvido no psicológico.

Em vez de privilegiar um lado da relação de identidade e tentar reduzir a mente ao cérebro, ou o cérebro à mente, pareceria que devemos reconhecer que se os eventos mentais são processos mentais, então devem ter propriedades físicas que os eventos cerebrais possuem. Mas, igualmente, temos de aceitar, de acordo com a simetria da identidade, que eventos cerebrais têm as propriedades mentais em virtude das quais os eventos mentais aos quais são idênticos são os tipos de eventos que eles são.

\*N. de T.: *O* *físicalismo*.

Qual deveria ser a nossa resposta a essas alturas? Bem, de uma coisa podemos ter certeza no caso do Super-Homem/Clark Kent, e no exemplo da água/H<sub>2</sub>O: é que não temos dois homens em nossas mãos, Clark Kent e o Super-Homem, e os dois tipos de coisas, água e H<sub>2</sub>O, mas somente um único item em ambos os casos. No caso da mente/corpo, parece que devemos estar preparados para aceitar que não há duas distintas correntes de eventos, eventos não-físicos ocorrendo na alma e eventos físicos ocorrendo no cérebro, mas somente um conjunto de eventos. Deste modo, podemos evitar um dualismo de substâncias, cada uma com suas naturezas disparatadas. Veja a questão no Exercício 3.2 antes de prosseguir na sua leitura.

### **EXERCÍCIO 3.2**

Dada a simetria da identidade e sua neutralidade no que diz respeito a ambas as metades da afirmação de identidade, que corrente de eventos deveríamos aceitar em nossa ontologia: físicos ou mentais? Quais são as razões para a sua escolha?

A evolução nos dá uma razão para privilegiarmos o físico. Parece razoável supor que, quando da formação inicial do universo, não havia mentes por lá, mas somente partículas rodopiando no espaço. Depois que os planetas se formaram e os fornos nucleares do Sol foram acesos, a vida surgiu de forma gradual. Eventualmente, depois que certo nível de complexidade física foi atingido, a mentalidade fez a sua aparição. Como ela surgiu não sabemos ainda, mas a partir do fato de que agora me dirijo a você, sabemos que, de alguma forma, ela surgiu. Também sabemos que o funcionamento mental está intimamente ligado ao que acontece no cérebro (veja a seção “A teoria da identi-

dade tipo-tipo e o reducionismo”). Tudo isso sugere enfaticamente que, ao contrário do que o dualismo afirma, há uma dependência unilateral do mental sobre o físico, que o físico, e não o mental, é que existe basicamente. Em consequência disso, parece não ser desprovido de razão privilegiar o físico em detrimento do mental, para concluir que, afinal de contas, em contraste com o dualismo, há somente substâncias físicas e não substâncias espirituais imateriais.

No entanto, mesmo se podemos desistir de postular dois tipos diferentes de substâncias ou eventos, mentais e físicos, parece que não podemos evitar a postulação de um dualismo de propriedades. Assim como suas propriedades físicas, aqueles estados cerebrais que são idênticos a estados mentais devem, ao que parece, também ter propriedades mentais.

Outro modo de ver isso é por meio da reflexão de que nem todos os estados cerebrais são idênticos a estados mentais. Deve haver estados cerebrais que não têm absolutamente nada a ver com a consciência – aqueles estados, por exemplo, que governam o funcionamento do sistema nervoso autônomo. Esses estados cerebrais possuirão somente propriedades físicas. No entanto, aqueles que são idênticos a estados mentais conscientes devem possuir uma outra camada de propriedades, precisamente mentais, sem o que a identidade desses processos cerebrais com processos mentais não poderia ser afirmada em primeiro lugar. O que está surgindo é que, dê você o colorido que quiser, podemos escapar de um dualismo de substâncias, mas não de um dualismo de propriedades. Uma mudança na vida mental de uma pessoa envolverá fundamentalmente mudanças físicas ou eventos no cérebro da pessoa e não eventos em uma alma não-física. No entanto, esses eventos físicos possuirão dois tipos

distintos de propriedades, físicos e mentais. O Quadro 3.1 resume essa posição e a contrasta com o dualismo, ilustrado do Quadro 3.2.

### **Qualia, acesso privilegiado e a irreduzibilidade do mental**

Outro argumento correlato que dá suporte à posição delineada na seção precedente tem a seguinte formulação: ao se fazer uma afirmação de identidade, deve haver alguns traços mediante os quais os itens que flanqueiam o sinal de identidade possam ser identificados independentemente uns dos outros, mediante os quais ambas as metades da identidade possam ser pinçadas, por assim dizer. Consequentemente, Clark Kent é identificado por meio de seus óculos de coruja, sua aparência nada excitante e seu modo desajeitado. O Super-Homem é identificado por sua visão perfeita, seu uniforme vistoso e pelo modo como acelera pelo ar por seus próprios meios. Os traços pelos quais Clark Kent é identificado não podem ser identificados com os traços pelos quais o Super-Homem é identificado. A miopia de

Clark Kent é uma propriedade diferente da visão perfeita do Super-Homem.

Do mesmo modo, os traços pelos quais os estados mentais de alguém são reconhecidos são diferentes dos traços pelos quais os processos cerebrais de alguém são discernidos. Eu sei que tenho uma dor de dente devido ao torturante latejar no meu dente. A partir da análise da fotografia de raios X do meu dente, sei também que ele está cariado e seu nervo danificado. O torturante latejar em meu dente é um traço bastante distinto dos processos neurofisiológicos revelados pela investigação científica. Eu poderia notar o latejar sem notar os processos físicos (caso não tivesse acesso a um aparelho de raios X), ou notar os processos sem notar um latejar (caso eu tivesse tomado um anestésico, ou tivesse sido hipnotizado, por exemplo). Ao que isso aponta fortemente é a que o latejar não pode ser identificado com as propriedades físicas do nervo danificado.

Os cientistas poderiam escanear o meu cérebro tanto quanto quisessem, mas não descobririam, meramente olhando, qual é a sensação que tenho ao sentir

**Quadro 3.1 A teoria mente/cérebro**

|  | E1                    | E2                  |                            |                     |
|--|-----------------------|---------------------|----------------------------|---------------------|
| Um conjunto de de eventos (E) aos quais se pode atribuir tanto descrições mentais quanto físicas | Pus o gato para fora? | fibras C disparando | Trouxe o gato para dentro? | fibras Z disparando |
|  | Propriedade mental    | Propriedade física  | Propriedade mental         | Propriedade física  |

**Quadro 3.2 Dualismo**

|                            |    |    |    |    |
|----------------------------|----|----|----|----|
| Eventos mentais na alma    | Mw | Mx | My | Mz |
| Eventos físicos no cérebro | Fa | Fb | Fc | Fd |

uma dor de dente. Dito de outra forma, eles não teriam acesso ao qualia da minha experiência, que, como salientado anteriormente (no Capítulo 1), estão, em princípio, disponíveis somente para mim. Igualmente, se um filme fosse feito da atividade do meu próprio cérebro e posteriormente esse filme fosse mostrado para mim, e nesse intervalo eu tivesse esquecido tudo sobre a ocasião em que fora feito, nem mesmo eu próprio poderia dizer, a partir simplesmente da observação de meu próprio cérebro, sobre o que estava pensando naquela ocasião.

Isso ajuda a reafirmar o ponto suscitado antes no Capítulo 1, qual seja, o de que, embora os cientistas e eu tenhamos igual acesso, em princípio, ao que acontece em nossos cérebros, somente eu tenho acesso aos conteúdos da minha mente, sem a mediação de qualquer processo de observação ou inferência baseada na evidência mais básica. O que esse acesso direto e imediato revela para mim não pode, ao que parece, ser meramente processos físicos cerebrais. Caso contrário, pela mera observação dos meus próprios processos cerebrais no filme, eu automaticamente saberia sobre o que estava pensando na ocasião em pauta. Sobre o que tenho consciência, quando tenho consciência de meus próprios pensamentos e sensações, deve, ao que parece, consistir em traços distintos das propriedades físicas públicas dos processos cerebrais.

Em uma conferência pública ministrada em Londres, Thomas Nagel expressou a piedosa esperança de que um dia um cientista do cérebro poderia olhar em seu cérebro e relatar algum processo que nele estivesse ocorrendo: "Veja! Eis como se apresenta neurologicamente, para Nagel, saborear o gosto de seu charuto!". No entanto, Nagel imediatamente acrescentou que, por enquanto, não temos a menor ideia de como, sentido para ele

subjetivamente e privadamente, o gosto de um charuto poderia ser simultaneamente um processo físico no cérebro, que em teoria pudesse ser observado por todos. Isso lembra a observação de Leibniz, citada anteriormente na "Introdução" do Capítulo 1, de que se o cérebro pudesse ser ampliado ao tamanho de um moinho, e fôssemos entrar nele, observaríamos processos mecânicos e nada que pudesse explicar a mentalidade e a consciência.

### **Estados mentais intencionais**

Até agora, o argumento contra a teoria da identidade tem sido o da não-reduzibilidade das qualidades privadas, subjetivas – os qualias – das sensações. No entanto, de longe, a maior parte do mental compõe-se de estados intencionais que representam, ou são dirigidos a, estados de coisas que não eles próprios, incluindo aqueles que não existem e talvez jamais tivessem existido. Se estados intencionais são identificados com processos cerebrais, então a questão surge instantaneamente: como podem estados físicos do cérebro ser sobre outros estados de coisas, inclusive aquelas que não existem? O argumento contra a identificação do intencional com o físico pode ser formalmente expresso do seguinte modo:

1. Os estados mentais intencionais representam, ou concernem, estados de coisas externas a si próprios, incluindo estados de coisas que não existem – e talvez jamais existissem. Por exemplo, crenças, desejos, sonhos, intenções, emoções, esperanças e pensamentos, assim como muitos outros estados mentais, possuem um conteúdo representacional.
2. Nenhum estado cerebral possui um conteúdo representacional.

3. Portanto, estados mentais intencionais não podem ser idênticos a estados cerebrais.

Agora, veja o Exercício 3.3 antes de prosseguir a sua leitura.

### **EXERCÍCIO 3.3**

O argumento esboçado é válido. Portanto, caso se resista a aceitar a conclusão, qualquer uma ou ambas as premissas devem ser rejeitadas. Qual premissa você pensa que deveria sair, se é que alguma delas deveria, e por quê?

O argumento é válido, e a primeira premissa é inquestionável. Consequentemente, caso desejemos rejeitar a conclusão, teremos de encontrar algum meio de rejeitar a segunda premissa. O pensamento que está por trás da segunda premissa e faz com que ela pareça altamente plausível é este: como é possível que arranjos de partículas puramente físicas no cérebro devam representar outros estados de coisas? Como podem coleções de átomos no vazio, por assim dizer, como Searle afirma em *Minds, Brains and Science*\*, concernir ou representar coisas que não sejam elas próprias, incluindo estados de coisas não-existentes?<sup>4</sup>

Em resposta a esse desafio, Jerry Fodor, um filósofo norte-americano, propôs a criação de uma abordagem puramente naturalista da intencionalidade. Tal abordagem se baseia em teorias científicas e procura dar uma explicação da intencionalidade em termos de entidades, propriedades, teorias e conceitos extraídos

das ciências físicas, tais como a química, a biologia e a física. A explanação tem como objetivo especificar as condições necessárias e suficientes em termos puramente físicos, para que a intencionalidade seja possível. É importante salientar que essas condições não serão condições *logicamente* necessárias e suficientes, como aquelas utilizadas em uma tentativa de fornecer uma definição formal de um conceito, por exemplo, o conceito de música.

Consequentemente, em sintonia com essa empreitada, diz Fodor:

Suponho que mais cedo ou mais tarde os fisicalistas concluirão o catálogo que têm estado compilando das derradeiras e irreduzíveis propriedades das coisas. Quando o fizerem, *rotação*, *atrativo* e *carga* (*propriedades microfísicas*) surgirão na lista deles. No entanto, *sobredade*\*\*, certamente, não: a intencionalidade simplesmente não vai a essa profundidade.<sup>5</sup>

Fodor, consequentemente, se opõe a filósofos tais como John Searle, Galen Strawson e Dale Jacquette,<sup>6</sup> que são cépticos quanto às reivindicações do naturalismo. Searle, por exemplo, sustenta que a intencionalidade é uma propriedade do cérebro, básica, emergente e irreduzível, que não pode ser reduzida a estados físicos ou propriedades, enquanto Jacquette vê a intencionalidade como um conceito explicativamente primitivo e indefinido:

Uma propriedade da mente, que não pode ser eliminada ou reduzida e mecanicamente replicável, tornando-a uma

\*N. de T.: Em português: Mentes, cérebros e ciência.

\*\*N. de T.: O termo *aboutness* é um neologismo cunhado a partir de uma preposição (*about*) mais um sufixo substantivo (*ness*). Seguindo mesmo processo teríamos em português o termo "sobreidade", composto de "sobre", no sentido de "a respeito de", mais o sufixo "dade". Nas ciências da linguagem *aboutness* é usado como equivalente de "assunto", "topicalidade". É o que responde à pergunta: sobre o que é este documento? Em filosofia é usado no sentido de intencionalidade original, ou ainda no de intencionalidade como faculdade de representação. No entanto, intencionalidade tem um sentido que não se coaduna com o de *aboutness*. Traduzir *aboutness* por "ser sobre" também não parece adequado, por gerar confusão com o "ser", usado substantivamente. Propomos a adoção de "sobreidade".

categoria nova de entidade no mundo material. A mente surge naturalmente a partir de matéria viva em sistemas biológicos complexos, em um estágio evolutivo comparativamente tardio. No entanto, devido à sua intencionalidade, a mente é qualitativamente diferente de coisas não-mentais, puramente mecânicas.<sup>7</sup>

Fodor discorda inteiramente, sustentando que deve haver um modo de explicar como os estados de um sistema físico, como o do cérebro, pode dizer respeito ou representar outros estados de coisas. Ele construirá, em contradição a Searle e Jacquette, uma teoria mecanicista em sua essência de como a intencionalidade é possível, baseado na crença de que o cérebro tem sua própria linguagem. Essa linguagem é, por hipótese, a linguagem do pensamento, que Fodor batiza de "mentalês". Não é semelhante a qualquer língua natural falada por povos do mundo; os símbolos do mentalês não são compostos por palavras, mas por estados computacionais do cérebro, realizados em processos sinápticos e neuronais. Esses processos, supõe-se, são combinados de modo análogo a palavras para gerar afirmações ou assertivas em mentalês, mas isso ainda deixa o problema de como tais estados supostamente possuem um conteúdo semântico ou significado.

Para responder a essa dificuldade, dois tipos de teorias naturalistas foram propostas. De acordo com a primeira teoria, uma representação, R, representa um estado de coisas, S, se e somente se R se parece com S. Isso é uma reminiscência da teoria pictórica do significado de Wittgenstein, que pode ser encontrada em seu *Tractatus-Logico-Philosophicus*<sup>\*</sup>: proposições elementares ou atômicas, isto é, proposições que não contêm algum

conectivo funcional de verdade, retratam ou representam possíveis estados de coisas. Eles fazem isso porque são compostos de nomes concatenados em uma estrutura particular, e o significado desses nomes são objetos simples. Quando a estrutura da proposição corresponde a uma estrutura de objetos que é real, em vez de meramente possível, a proposição é verdadeira; caso contrário, é falsa. Wittgenstein nunca fornece exemplos de nomes ou objetos na formulação desta teoria, mas é quase certo que não devem ser compreendidos segundo o modelo do uso comum de todos os dias desses termos. A inspiração para a teoria foi a utilização, em uma delegacia parisiense, de modelos para representar um acidente de trânsito, mostrando as posições relativas dos veículos e pedestres envolvidos. É interessante que, em um de seus primeiros cadernos de notas, Wittgenstein escreveu: "Pois o pensamento é um tipo de linguagem. Pois o pensamento é também, certamente, um retrato lógico da proposição, e, portanto, é somente um tipo de proposição".<sup>8</sup> E, em uma carta a Bertrand Russell, Wittgenstein ampliou esse comentário, mantendo que um pensamento "não consiste em palavras, mas em componentes psíquicos que correspondem a palavras e têm o mesmo tipo de relação para com a realidade que as palavras."<sup>9</sup> Se a palavra "físico" é substituída por "psíquico", então a similaridade entre a teoria pictórica do significado de Wittgenstein e a teoria da parecência naturalista do significado, que estamos analisando, torna-se aparente.

Infelizmente, no entanto, a mera parecência não será suficiente para a representação. Felipe e Pedro são gêmeos idênticos. No entanto, nenhum deles representa o outro. Nem, no entanto, a parecência é necessária para a representação:

\*N. de T.: Traduzido para o português com o mesmo título. São Paulo: EDUSP, 3. ed., 2001.

o nome "Pedro" não tem nada a ver com o próprio Pedro, embora possa ser usado para referir a ele. Além disso, enquanto a parecença é uma noção simétrica – se Pedro se parece com Felipe, então Felipe igualmente se parece com Pedro –, a representação não o é. Uma fotografia de passaporte pode representar o portador do passaporte, mas o portador não representa a fotografia.

A segunda teoria tenta explicar a *sobredade* em termos de causalidade, especificamente, de covariação causal. Para explicar o que isso significa, temos o exemplo seguinte. Conhecemos o tipo de anel que uma árvore deposita em um determinado ano, sua espessura e densidade, por exemplo, depende de como foi o clima naquele ano. A correlação é facilmente estabelecida por meio do corte da árvore e da comparação da natureza de seus anéis com os registros metereológicos dos diferentes anos. Deste modo, somos capazes de futuramente, sem consultar registros, estabelecer como o clima deveria ter sido em anos específicos, independentemente da manutenção de registros. Os anéis da árvore são, em grande parte, indicadores confiáveis do clima em anos passados. Há, portanto, um sentido de que os anéis das árvores contêm informação sobre o meio ambiente externo da árvore ao longo dos anos. Os anéis variam em um modo confiável e semelhante a uma lei de acordo com mudanças no meio ambiente da árvore; isto é, eles covariam causalmente. De um modo semelhante, os estados internos de um termostato variam como resultado de mudanças na temperatura ambiente. De fato, construímos termostatos com esse propósito ou função específica, a fim de que interruptores possam ser ativados quando o estado interno atinge um determinado valor, para ligar ou desligar o sistema de aquecimento. A temperatura ambiente causa nos estados internos

do termostato uma variação sistemática e confiável, e, portanto, de acordo com a análise da intencionalidade, em termos de covariação causal, pode-se dizer que eles representam aquela temperatura. Marvin Minsky, um cientista e pesquisador no Michigan Institute of Technology, quer levar essa questão ainda mais adiante, sustentando, com aparente seriedade, que este termostato tem três crenças: está muito quente aqui; está muito frio aqui; aqui está perfeitamente bem.

No entanto, não necessitamos atribuir a termostatos ou a anéis de árvores a literal manutenção de crenças sobre o meio ambiente. Para as próprias árvores ou termostatos, seus estados internos podem ser desprovidos de significado, visto que nem árvores nem termostatos têm consciência. A um termostato, desprovido de consciência, não se pode literalmente atribuir crenças falsas ou verdadeiras sobre o meio ambiente. Uma representação somente conta como uma representação se pode representar ou de fato representa algo para um ser consciente. Consequentemente, nuvens, por si mesmas, não representam baleias ou camelos, independentemente de quanto possam parecer-se com uma baleia ou um camelo. Mas adotar este ponto de vista pressupõe uma perspectiva consciente do mundo, algo de que árvores e termostatos são notoriamente incapazes.

Os exemplos de covariação causal recém-descritos são todos instâncias do que é conhecido como indicação confiável. O estado interno do termostato, variando como uma lei de acordo com a temperatura ambiente que o faz mudar, indica confiavelmente aquela temperatura; os anéis das árvores, de forma semelhante, confiavelmente indicam como foi o clima em um dado ano; e nuvens pretas indicam confiavelmente a vinda de uma tormenta. Desse modo somos conduzidos

a uma teoria do significado natural: aquelas nuvens pretas *significam* uma tormenta, aqueles anéis de árvore *significam* que o clima foi muito seco naqueles anos, e assim por diante. Aplicada ao mentalês a hipótese da linguagem do pensamento, a afirmação é de que os estados computacionais do cérebro são causados de modos regulares e confiáveis que covariam com o ambiente externo, então aqueles estados são sobre aqueles estados externos de coisas e os podem “significar”. De forma mais precisa: um estado cerebral, B, representa um estado externo de coisas, A, se e somente se B confiavelmente indica A. Isso afirma as condições causalmente, não logicamente necessárias e suficientes para que B represente A, ou seja, sobre ele. Falar sobre representação ou verdade não é, quanto ao significado, equivalente a falar sobre indicação confiável: a redução proposta não é uma redução analítica, mas naturalista ontológica. Negar que, em última análise, a intencionalidade resulta em indicação confiável não é uma contradição em termos, assim como não é contraditório negar que água seja  $H_2O$ . Negar que água seja  $H_2O$  é simplesmente cometer um engano factual, e, semelhantemente, negar que a intencionalidade possa ser reduzida à indicação confiável simplesmente acontece de ser factualmente errado. Ou nisso é o que os defensores da indicação confiável gostariam que acreditássemos.

No entanto, a teoria da indicação confiável deve se deparar com uma série de dificuldades sérias. Primeiro, é necessário que B seja confiavelmente causado por A para que B represente A? Caso afirmativo, então, se B não é causado por A, ou ao menos não é confiavelmente causado por A, então B não pode representar A. O problema é que eu posso pensar a respeito de uma quantidade de coisas que não existem e, portanto, não posso ter causado

meus pensamentos a respeito delas. Por exemplo, eu posso pensar sobre a Fonte da Juventude, sobre Pégaso, sobre uma montanha de ouro, muito embora eu não tenha nenhuma experiência dessas coisas, e posso crer que fadas estão no fundo do meu quintal, muito embora não haja fadas a serem encontradas lá. Os que defendem a teoria da indicação confiável podem, no entanto, ter um meio de contornar essa dificuldade. Os componentes desse pensamento, eles poderiam sustentar, deveriam ser causados de modos regulares e confiáveis. Consequentemente, o pensamento sobre o ouro foi causado de modos regulares pela visão de amostras dele, e semelhantemente com o pensamento sobre uma montanha. Trata-se meramente que essas duas coisas são postas juntas na imaginação e, embora não haja montanha de ouro alguma que corresponda a ela na realidade, os componentes realmente têm contrapartes externas que os suscitam de um modo confiável. A similaridade dessa abordagem com a teoria de Hume sobre a aquisição de conceitos deve ser aparente, pois Hume afirmava que todos os conceitos, sem exceção, eram cópias mais pálidas de impressões sensoriais. Em outras palavras, todos os materiais do pensamento tiveram sua origem em impressões sensoriais. Podemos então usar de nossa imaginação para formar conceitos de coisas que nunca vimos, tais como montanhas de ouro ou centauros, por meio da união de conceitos simples originalmente derivados de impressões sensoriais para, assim, formar conceitos mais complexos.

Talvez, então, o defensor da teoria da indicação confiável possa, ao final de contas, escapar da objeção de que podemos pensar sobre coisas das quais não temos nenhuma experiência direta. No entanto, a teoria não está isenta de riscos ainda. A análise ainda afirma que para que B represente A, uma razão suficiente é que B

seja confiavelmente causado por B. Mas veja agora o seguinte exemplo: meu escâner adquiriu uma falha sistemática; então, regular e sistematicamente ele não escaneará a palavra "não". Então, quando eu escaneio os Dez Mandamentos, posso estar seguro todas as vezes que a cópia que faço me adjura a cometer assassinato, adultério e roubo. Mas não se pode dizer que minha cópia representa os Dez Mandamentos; de forma muito clara, ela seriamente distorce a representação deles. Não obstante, se a indicação confiável é suficiente para que a representação ocorra, teríamos de afirmar o contrário, ou seja, que minha cópia dos Dez Mandamentos com todos os seus "nãos" deixados de fora efetivamente representa os Dez Mandamentos, afinal de contas. No entanto, como a cópia simplesmente não representa os Dez Mandamentos, pareceria que a indicação confiável não pode ser uma condição causal suficiente para que B represente A. Certamente, se as cópias fossem criadas do modo correto, representariam os Dez Mandamentos, mas como devemos especificar qual seja o modo correto? Mergulharemos imediatamente em circularidade, caso insistamos que o meio correto é aquele que conduz a uma produção de representações acuradas.

Finalmente, há mais três objeções à análise da covariação causal. Primeiro, a covariação causal é um conceito transitivo. Suponha que a cadeia causal, que vem do estado de coisas externo, A, para o estado cerebral, B, tem mais do que uma conexão, então entre A e B há mais duas conexões, C e D, que consistem em eventos no sistema nervoso, por exemplo, eventos neurais surgindo a partir de estimulação da retina por meio de A. Consequentemente, a cada vez que B covariar com A, também covariará com C e D. Não obstante, a percepção ou pensamento dirá respeito a A, e não a C ou D.

Segundo, na mesma esteira, há o problema apresentado pela coextensão. O conceito de um triângulo é diferente, quanto ao seu significado, do conceito de um trilateral; no entanto, cada triângulo que vejo deve também ser necessariamente um trilateral. Os conceitos são diferentes quanto aos seus significados. Contudo, o escopo de objetos aos quais se aplicam, suas extensões, deve coincidir logicamente. Porém, quando vejo certa figura diante de mim em alguma ocasião, formo a crença de que vejo um triângulo diante de mim e não a crença de que vejo um trilateral. Mas como se pode explicar a crença de que eu vejo um triângulo e não um trilateral, dado o fato de que as extensões dos triângulos e dos trilaterais são necessariamente coextensas? Como pode a covariação causal explicar a diferença entre as duas crenças que eu poderia adquirir nessa situação?

Finalmente, um problema se coloca pela noção de coconstância, a essa relacionada. Quine imagina um falante nativo de uma língua que pronuncia "Gavagai" no exato momento em que um coelho passa. Ele quer dizer "Aquilo é um coelho" ou "Aquilo é uma parte não separada de um coelho"? O conceito de coelho e o conceito de uma parte não separada de um coelho são diferentes quanto aos seus significados, e, diferentemente dos conceitos de triângulos e trilaterais, os conceitos não são coextensos. No entanto, sempre que o conceito de coelho for instanciado, então necessariamente o conceito de uma parte não separada de coelho também será. Mas qual conceito adquire o falante nativo – a crença de que lá vai um coelho, ou a crença de que lá vai uma parte não separada de um coelho? Assim como no caso da coextensão, como pode o defensor da teoria causal explicar a diferença entre as duas crenças? As prospecções de uma abordagem

naturalista da intencionalidade em termos de indicação confiável e covariação causal, portanto, não parecem brilhantes. No entanto, a proposta ainda tem seus leais defensores.

Antes de concluirmos, há um defeito em comum com o qual tanto a teoria da representação quanto a da covariação causal se deparam, qual seja, o de que não podem explicar a normatividade do significado. O que isto significa é que, porque as expressões adquirem seus significados e uso por convenção, devem ter um uso correto. Portanto, também devem ter um uso incorreto, e os falantes da língua devem distingui-los e, caso seja necessário, corrigir um uso incorreto. Mas as teorias da representação e da covariação causal não fazem menção alguma sobre o uso correto e incorreto de palavras por usuários da língua. Portanto, a normatividade do significado fica não reconhecida e não explicada. Retornaremos e aprofundaremos a discussão sobre essa questão no Capítulo 8, seção “A falácia mereológica”, em um contexto de uma análise da contribuição de Aristóteles à filosofia da mente.

### **Racionalidade e normatividade**

Há um argumento ainda mais forte contra a possibilidade de ser dada uma explicação fisicalista da intencionalidade. Será retomado do Capítulo 1 que, dado o fato de que tais estados podem ser logicamente relacionados uns aos outros, estão sujeitos às restrições da rationalidade e da normatividade. Crenças, por exemplo, podem conflitar, e caso nos demos conta de um conflito entre crenças, então rationalmente devemos abandonar uma delas ou ambas. Exigimos rationalidade tanto de nós mesmos quanto dos outros, e tem sido defendido pelo filósofo Donald Davidson que, ao interpretar o comportamento dos outros, fazemos isso tentando enquadrar

os estados intencionais que atribuímos a essa pessoa – suas intenções, desejos, crenças e assim por diante – em um todo coerente. Davidson concorda que as pessoas possam ser irrationais, mas a irrationalidade tem sua existência sob a dependência de um pano de fundo geral de rationalidade. “Imaginar um animal totalmente irracional”, diz ele, “é imaginar um animal sem pensamentos”.<sup>10</sup>

Davidson desenvolve ainda mais essa linha de pensamento. Ele argumenta que o domínio do físico e o domínio do mental, onde ele tem estados intencionais em mente, estão sujeitos a uma gama de restrições completamente diferentes. Um exemplo de tal restrição na teoria da física é dado pelo princípio da transitividade para o comprimento: se A é mais comprido do que B, e B é mais comprido do que C, então A deve ser mais comprido do que C. Caso um princípio desses seja abandonado, então a coerência da teoria física está ameaçada. Assim sendo, a transitividade do comprimento é um elemento constitutivo da teoria-física. O elemento constitutivo da contraparte, em relação ao domínio mental, é o que Davidson chama “o ideal constitutivo da rationalidade”, uma ideia que encontramos um instante atrás ao falar da necessidade de atribuir um padrão de estados intencionais razoavelmente coerentes e consistentes a outras pessoas, sob pena de perder a oportunidade de pensá-las como pessoas, e de tratá-las devidamente.

Devido ao fato de o mental e o físico serem governados por tipos de constituintes de tipos tão radicalmente diferentes, Davidson acredita que nenhum enquadramento em larga escala dos estados físicos e mentais será possível. Assim, leis psicofísicas conectando eventos sob suas descrições mentais com suas descrições físicas serão descartadas. Em outras palavras, leis de ponte bicondicionalis, do tipo

mencionado antes na seção “A teoria da identidade tipo-tipo e o reducionismo”, não serão vindouras, e com elas qualquer prospecto de redução do mental a traços físicos, como a teoria da identidade tipo-tipo requer. Por essa razão, o conceito de holismo de estados intencionais e o “ideal constitutivo de racionalidade”, que governa sua atribuição, constituem ainda outro argumento contra a teoria da identidade tipo-tipo.

### *A identidade espécime-espécime e o prospecto de redução*

Salientamos anteriormente que a teoria espécime-espécime parecia ser a alternativa mais plausível à teoria tipo-tipo. No entanto, em acréscimo às objeções que se aplicam às teorias da identidade em geral, a teoria espécime-espécime tem de se deparar com alguns problemas que lhes são próprios.

O primeiro problema que a teoria tem de enfrentar é o de que, embora cada espécime de evento que tem uma propriedade mental tenha também uma propriedade física, não haverá relação sistemática entre o tipo de evento mental que um espécime instancia e o tipo de evento físico que ele também exemplifica. Para que se compreenda isso mais claramente, veja as propriedades da forma e da cor. Cada objeto colorido terá uma forma, e, igualmente, pode-se dizer, cada item com uma forma deve possuir uma cor (devo ignorar a possibilidade de objetos perfeitamente transparentes ou invisíveis). No entanto, não haverá relação sistemática alguma entre as propriedades de cor e de forma, tal que um objeto possa ser, digamos, retangular se e somente se for vermelho; ou que um objeto possa ser azul se e somente se for redondo. Cores e formas podem variar com muita independência uma da outra, e sua rela-

ção é bastante acidental. Portanto, não há prospecto de explicar cores em termos de formas ou formas em termos de cores. O mesmo pareceria ser verdade do fysicalismo espécime-espécime como ele tem sido caracterizado até agora. Não haverá qualquer modo sistemático de explicar por que eventos particulares coinstanciam conceitos de um determinado tipo mental e de um determinado tipo físico. A existência de um traço mental será tão aparentemente não-relacionada à existência do traço físico, como são a cor e a forma de um objeto um para o outro. Eis por que a tese da identidade espécime-espécime tem sido estigmatizada por filósofos como não pertencendo muito a um fysicalismo: porque ela não permite que estados físicos desempenhem algum papel explanatório na abordagem da presença de estados mentais. A coexistência de traços mentais e físicos deve ser aceita como um fato bruto, inexplicável, consequentemente fazendo com que a teoria espécime-espécime da identidade seja difícil de abraçar.<sup>11</sup>

No entanto, há razão para se acreditar que essa dificuldade particular possa ser superada por meio da invocação do que é conhecido como a hipótese da superveniência. Uma caracterização completa desta proposta deve esperar até o Capítulo 6, mas seu ataque principal é que, para que possam existir, traços mentais dependem de traços físicos, de modo tal que não pode haver mudança mental a menos que haja uma mudança física. O contrário não se aplica, porque espécimes de tipos diferentes de estado físico poderiam instanciar espécimes de um e mesmo tipo de estado mental. Isso é um outro modo de dizer o ponto anteriormente levantado, o de que estados mentais são multiplamente realizáveis em diferentes tipos de estados físicos. A superveniência, caso possa ser posta em funcionamento,

tem ao menos a virtude de trazer o mental e o físico em uma relação mais próxima um com o outro, e nessa medida pode superar o estigma que salientamos ter sido imputado à teoria. No entanto, a superveniência e seus problemas adjacentes deverão esperar até um capítulo posterior.

### *A identidade espécime-espécime e a impossibilidade de redução*

A teoria espécime-espécime está ainda isenta de riscos, no entanto. Há um poderoso argumento contra ela, no sentido de que a possibilidade de realizações múltiplas dos estados mentais, que a teoria espécime-espécime enaltece, proíbe toda a redução do mental ao físico. Para que possamos compreender por que isto deve assim ser, devemos lembrar que a redução de um conjunto de fenômenos a outro parece requerer a possibilidade da formulação de leis de ponte bicondicionais. Uma lei desse tipo, será lembrado na seção "A teoria da identidade tipo-tipo e o reducionismo", requer que, para que um estado mental  $M$  seja redutível a um estado físico  $F$  e a ele identificável, deva ser o caso de que  $M$  ocorra se e somente se  $F$  ocorrer. A possibilidade de realização múltipla parece quebrar este requisito. Isso porque  $M$  pode ser realizado em uma variedade indefinida de estados físicos. Em outras palavras, um único estado físico  $F$  não é necessário para a ocorrência de  $M$ :  $F_1, F_2, F_3 \dots$  ou  $F_n$  é suficiente para que  $M$  ocorra, mas não necessário. No entanto, a lei da ponte requer que qualquer estado físico candidato ao qual  $M$  possa ser reduzido deva ser tanto necessário quanto suficiente para a ocorrência de  $M$ . Agora faça uma pausa para examinar o Exercício 3.4 antes de prosseguir sua leitura.

### **EXERCÍCIO 3.4**

Você pode ver uma possível resposta ao desafio delineado? Pense a esse respeito e discuta com outra pessoa.

A resposta do reducionista ao desafio é sustentar que uma propriedade física pode ser encontrada, ou melhor, construída, que pode servir como condição necessária e suficiente para  $M$ . Tudo o que necessitamos fazer é construir a propriedade disjuntiva ou  $F$  ou  $F_1$  ou  $F_2$  ou  $F_3 \dots$  ou  $F_n$ .

No entanto, há um problema com esta estratégia. Veja o Exercício 3.5 antes de prosseguir.

### **EXERCÍCIO 3.5**

Você pode ver algum problema na resposta do reducionista? Dica: pense sobre o que são os traços de propriedades, e determine se as assim chamadas propriedades disjuntivas podem qualificar-se.

O problema com a resposta do reducionista é a seguinte: mesmo que  $F, F_1, F_2$  e  $F_3$ , etc. sejam propriedades quando cada uma é considerada singularmente por si própria, não se segue que a combinação disjuntiva delas represente uma propriedade genuína. As propriedades são comumente pensadas como pontos marcantes de parecência, similaridades compartilhadas, entre itens particulares. No entanto,  $F_1, F_2, F_3$ , etc. poderiam ser estados físicos muito radicalmente diferentes, especialmente quando lembramos que tais estados não necessitam de ser restritos a organismos, mas poderiam inclusive englobar estados de máquinas tais como robôs, computadores ou andróides. Em última instância, a possibilidade de que estados mentais pudessem ser encarnados em arranjos distintos daqueles orga-

nismos baseados em carbono não pode ser descartada por antecipação. Mas, então, a disjunção conterá tal coleção heterogênea de itens que não terão similaridades significantes entre eles, e, assim, parece altamente artificial que se fale deles como constituindo uma única propriedade. Em acréscimo, não será certamente uma única teoria científica que poderá abarcar tal confusa coleção de propriedades.<sup>12</sup>

Outro problema que se apresenta para reducionistas é que, em princípio, parece que a análise disjuntiva nunca poderia ser completada, porque, dada sua própria natureza, seu fim está em aberto. Caso tentemos completar a análise, no entanto, por meio do acréscimo à série F ou F1 ou F2 ou F3, etc. da expressão "e qualquer outro estado físico que instancia M", então uma circularidade é importada para a análise, pela menção do termo mental M. Uma análise que contenha o próprio termo "mental" que se supõe que ela elimine, pela redução ao físico, claramente não pode funcionar. Se essas dificuldades para o reducionismo podem ser superadas em última instância, deixo ao encargo do leitor prosseguir a reflexão.

Uma terceira severa dificuldade para a teoria espécime-espécime foi apontada por John Searle em seu livro *The Rediscovery of the Mind\**.<sup>13</sup> De acordo com a teoria espécime-espécime, você e eu poderíamos estar no mesmo estado mental – por exemplo, pensando no que teremos para o jantar – e, no entanto, nossos estados neurológicos poderiam ser diferentes um do outro, devido à possibilidade de realização múltipla do mental pelo físico. Mas o que é então que faz de nossos distintos estados mentais instanciações de um e mesmo padrão de pensamento? Veja o Exercício 3.6 antes de ver minha resposta a esta questão.

### **EXERCÍCIO 3.6**

Qual seria a resposta dada pelo senso comum a essa pergunta, em sua opinião, e por que ela seria fatal ao materialismo redutivo? Qual resposta contrastante deve ser dada pelo materialista redutivo e por quê? Pense sobre isso sozinho, e, se possível, discuta com alguém mais. Depois compare com o meu comentário, a seguir.

A resposta do senso comum à pergunta apresentada no último parágrafo é que cada espécime de tipo diferente de estado neurofisiológico possui o mesmo tipo de propriedade mental, qual seja, a propriedade de pensar o que há para o jantar. Isso porque os diferentes espécimes são instanciações do mesmo tipo de estado mental. Parece-me que esta é a resposta correta à questão, e ajuda a constituir ainda outra razão para que não se subscreva o estrito fisicalismo. Claramente, no entanto, não satisfará o materialista redutivo. A razão é a de que traços mentais não podem ser citados para explicar por que espécimes de tipos diferentes de estados cerebrais são ainda incorporações de um e mesmo padrão de pensamento, porque o derradeiro objetivo do teórico da identidade é reduzir traços mentais a traços físicos, eliminar o mental, de fato, em favor do material. Consequentemente, como salienta Searle, para o materialista dedicado, algumas respostas não-mentais devem ser dadas, tais como a de que estados neurológicos distintos executam a mesma função na vida de um organismo. É justamente desse modo que o fisicalismo espécime-espécime e a possibilidade de realização múltipla conduzem à teoria mais recente da relação mente/corpo, qual seja, a do funcionalismo. No entan-

\*N. de T.: Em português: *A redescoberta da mente*.

to, devemos esperar até o Capítulo 5 para uma exposição completa e para a avaliação desta posição mais recente da relação entre o mental e o físico.

### O argumento de Kripke contra a teoria da identidade

Guardei para perto do fim deste capítulo o argumento mais complicado, e possivelmente o mais poderoso, contra a teoria da identidade mente/cérebro, que fala igualmente tanto contra a versão tipo-tipo quanto contra a espécime-espécime. O argumento que examinaremos deve-se a Saul Kripke, filósofo norte-americano, e encontra-se em seu livro *Naming and Necessity (A nomeação e a necessidade)*.<sup>14</sup>

Para que se compreenda a objeção de Kripke, precisamos em primeiro lugar lembrar a afirmação dos defensores da teoria da identidade, de que a identidade do mental e do físico é uma verdade contingente e não analítica. Tudo o que isso queria dizer era que não é possível dizer mediante o mero exame da afirmação relevante de identidade se ela era verdadeira ou não. A afirmação de que os estados mentais são idênticos a estados cerebrais, se verdadeira, não é verdadeira *a priori*, em virtude do significado dos termos que flanqueiam o sinal de identidade. Como vimos anteriormente, falar a respeito de estados mentais é diferente, quanto ao significado, de falar sobre estados cerebrais.

Kripke, no entanto, argumentou que se, por uma questão de fato contingente, estados mentais são estados cerebrais, então eles são essencialmente estados cerebrais. Isso não quer dizer que seja necessariamente verdade que a dor, por exemplo, seja física. A diferença entre o que os filósofos chamam de necessidade de *dicto* e de *re* é aqui vital. A necessidade de *dicto* é uma propriedade da linguagem,

e as verdades analíticas são uma exemplificação dela. A necessidade de *re* é uma propriedade do mundo. A dor é necessariamente física e a necessidade aqui é de *re*, não uma necessidade de *dicto*.

Para que se entenda por que Kripke faz a afirmação de *re* de que a dor é necessariamente ou essencialmente física, devemos primeiro ver a distinção que ele estabelece entre o que chama de designadores rígidos e não-rígidos. O próprio exemplo de Kripke de designador não-rígido é a descrição definida, "O primeiro Chefe de Correio Geral dos EUA". Essa expressão se refere a Benjamin Franklin, que foi, ao que parece, o primeiro Chefe de Correio Geral. Mas, simplesmente, ele pode não o ter sido, o que quer dizer que há um mundo possível no qual "O primeiro Chefe de Correio Geral dos EUA" não se refere a Benjamin Franklin, mas a uma outra pessoa. Em geral, todas as descrições definidas serão designadores não-rígidos e poderiam se referir a diferentes indivíduos em diferentes mundos possíveis.

Por outro lado, Kripke afirma que o nome próprio "Benjamin Franklin" (e os nomes próprios em geral) é um designador rígido, porque ele seleciona um e o mesmo indivíduo em todos os mundos possíveis – isto é, ele refere necessariamente a Benjamin Franklin, e é inconcebível que pudesse selecionar alguma outra pessoa.

Agora veja o nome "Lewis Carroll". Como um designador rígido, ele necessariamente seleciona o mesmo indivíduo em todos os mundos possíveis. O nome "Charles Dodgson" também identifica a mesma pessoa em todos os mundos possíveis. Agora seria possível que cada um desses nomes tivesse selecionado uma pessoa diferente da outra. Assim, "Lewis Carroll" poderia ter se referido a um indivíduo particular que existiu em todos os mundos possíveis, e "Charles Dodgson"

poderia ter se referido a um outro indivíduo *diferente*, que existiu em todos os mundos possíveis. No entanto, embora esses nomes pudessem ter se referido a pessoas diferentes, de fato não o fizeram, porque, como é bem conhecido, "Lewis Carroll" era o pseudônimo que Charles Dodgson usou quando compôs seus livros sobre Alice. Consequentemente, podemos formular a afirmação de identidade:

Lewis Carroll = Charles Dodgson

Cada um dos nomes que flanqueiam o sinal de identidade é um designador rígido; assim, cada nome necessariamente refere o mesmo indivíduo em todos os mundos possíveis. Os indivíduos assim referidos poderiam não ser numericamente idênticos, mas nesse caso eram, porque é verdade que Lewis Carroll era Charles Dodgson. Isso conduz a um resultado interessante: não há um mundo possível em que o nome "Lewis Carroll" identifique uma pessoa individual que não seja também identificada pelo nome "Charles Dodgson". Fazendo essa colocação negativamente, não há mundo possível, nenhum conjunto concebível de circunstâncias, em que a referência de "Lewis Carroll" pudesse ser diferente da referência a "Charles Dodgson". Em todos os mundos possíveis, esses dois nomes devem necessariamente se referir a uma e à mesma pessoa individual. Assim, Lewis Carroll não é meramente idêntico a Charles Dodgson, ele é necessariamente idêntico (no sentido de necessidade de *re*) a Charles Dodgson, porque não há circunstância concebível, nenhum mundo possível, em que essa identidade não se mantenha. Supor o contrário seria como supor que algo, neste caso a pessoa conhecida, quer como Lewis Carroll, quer como Charles Dodgson, pudesse deixar de ser

idêntica a si mesma, o que é simplesmente uma absurdade. Por outro lado, aquelas afirmações de identidade nas quais o sinal de identidade é flanqueado exclusivamente por designadores rígidos, uma afirmação de identidade da forma

Lewis Carroll = o autor de *Alice in Wonderland*\*

nunca pode ser possível em todos os mundos possíveis. Isso porque a expressão à direita do sinal de identidade, "O autor de *Alice in Wonderland*", é uma descrição definida. Isso quer dizer que ela é um designador não-rígido que pode se referir a diferentes pessoas em mundos possíveis diferentes. Essa afirmação é, certamente, verdadeira, mas não exibe necessidade de *re*, porque não se mantém ao longo do espectro de todos os mundos possíveis. Pode haver um mundo no qual o autor de *Alice in Wonderland* fosse alguém distinto de Lewis Carroll, digamos H. G. Wells, por exemplo.

Kripke estende a noção de designador rígido para cobrir não somente nomes próprios, mas termos que nomeiam espécies naturais, como ouro, cloro, água e  $H_2O$ . Consequentemente, "água" como um designador rígido seleciona o mesmo tipo de material em todos os mundos possíveis. O mesmo faz " $H_2O$ ". Mas "água" = " $H_2O$ " expressa uma identidade verdadeira, uma vez que "água" e " $H_2O$ ", ambos, se referem a um e mesmo material em todos os mundos possíveis. Assim, embora tivesse de ser descoberto empiricamente que água é  $H_2O$ , e assim foi conhecido somente com base na experiência, diferentemente das verdades analíticas da matemática e da geometria, que poderiam ser conhecidas independentemente da experiência, a afirmação de identidade, "água =  $H_2O$ ",

\*N. de T: Em português: *Alice no País das Maravilhas*.

bem como "Lewis Carroll = Charles Dodgson", exibe necessidade de *re*. Ou seja, não há mundo possível algum no qual "água = H<sub>2</sub>O" seja falso.

"Água = H<sub>2</sub>O" é uma das identidades científicas listadas anteriormente na seção "A teoria da identidade tipo-tipo". Você se lembrará de que, nesses tipos de afirmação, o termo à direita deve ser pensado como dando a natureza escondida do fenômeno referido à esquerda. De acordo com o defensor da teoria da identidade, o mesmo se mostra verdadeiro para a afirmação de identidade "Dor = fibras C disparando". Mas, como vimos, de acordo com Kripke essa identidade, assim como todas as demais identidades científicas, se for verdadeira, deve ser verdadeira em todos os mundos possíveis. Em outras palavras, não há mundo possível que possamos conceber em que a dor não seja idêntica a fibras C disparando, porque a identidade da dor com fibras C disparando é uma identidade necessária de *re*.

Mas é precisamente aqui que o defensor da teoria da identidade se vê liberado, de acordo com Kripke, porque podemos conceber duas possibilidades. Primeiro, é possível conceber uma situação em que uma pessoa tem uma dor, mas não há fibras C disparando. Tudo o que há de essencial para a dor é o modo como ela é sentida. O que possa ou não estar acontecendo neurologicamente é irrelevante, uma vez que tudo o que é requerido é o seguinte: se uma sensação lhe parece ser dolorida, então é dolorida. Há, portanto, um mundo possível em que a dor ocorre sem a presença de fibras C disparando. O defensor da teoria da identidade protestaria em vão: "Mas o que você sente não pode ser dor, porque suas fibras C não estão disparando". A situação é precisamente diferente daquela em que você vê um líquido sem cor em uma jarra e julga que seja água. Aqui você poderia estar

enganado, porque se a natureza interna da matéria na jarra for H<sub>2</sub>SO<sub>4</sub>, e não H<sub>2</sub>O, então o líquido é ácido sulfúrico e não água. O que faz com que uma dor doa, no entanto, não é uma natureza interna escondida, da qual você pode não ter conhecimento, mas simplesmente o modo como você a sente.

Na eventualidade de este ponto não estar suficientemente claro, permita-me tentar uma outra apresentação. Sabemos que a água é idêntica a H<sub>2</sub>O, e, portanto, de acordo com Kripke, ela é necessariamente (com necessidade de *re*) idêntica a H<sub>2</sub>O. Assim, não há mundo possível algum em que ela não seja H<sub>2</sub>O. No entanto, não podemos imaginar um mundo que contenha água que tenha a fórmula química XYZ? A resposta de Kripke é que, em um caso como esse, não se está efetivamente conseguindo imaginar água, de forma alguma, mas somente alguma coisa aquosa que se parece com água e se comporta como água, mas que, na realidade, não é água, porque sua composição interna é XYZ e não H<sub>2</sub>O. Dito de outra forma, contrariamente ao que se pensa, não se consegue imaginar a água separadamente de H<sub>2</sub>O, porque o que se está imaginando não é água, mas alguma outra coisa. Agora, a mesma estrutura se aplica no caso da dor, por exemplo? Posso imaginar efetivamente a ocorrência de uma dor, sem a ocorrência de fibras C disparando? O defensor da teoria da identidade, certamente, se vê obrigado a dizer que não posso, assim como não posso imaginar a água sem que H<sub>2</sub>O esteja presente. O que o defensor da teoria da identidade tem que manter é que, quando eu acredito que estou imaginando dor na ausência de fibras C disparando, estou somente imaginando algo que se parece com uma dor, e não uma dor propriamente, porque isso que se parece com uma dor não possui a constituição interna correta. Mas aqui é onde

Kripke deseja fazer o seu ponto essencial, qual seja, o de que o caso da água e o caso da dor são radicalmente diferentes. Tudo o que é necessário para que uma dor seja uma dor é que ela seja sentida como uma dor. Não há, nem pode haver, qualquer distinção entre uma dor e algo que se parece com uma dor do mesmo modo como há uma distinção entre a água e algo que se parece com água. Você poderia ter algo que se parecesse exatamente com água, mas na realidade não fosse água, porque a parecença não é essencial para que ela seja aquilo que é, enquanto sua constituição interna, por outro lado, é essencial. Mas você não poderia ter algo que se parecesse exatamente com uma dor, mas que não fosse uma dor, porque o que é essencial para que uma dor seja uma dor é simplesmente a maneira como ela é sentida. Assim, parece que podemos genuinamente imaginar dores ocorrendo na ausência de fibras C disparando, e isso quer dizer que dores não são idênticas a fibras C disparando. Do mesmo modo, podemos imaginar um mundo em que as fibras C desparem, mas não sejam acompanhadas de dor. Isso quer dizer que há mundos possíveis em que os mesmos tipos de coisas se passam no cérebro das criaturas, como ocorrem em nossos cérebros neste mundo, exceto que, diferentemente de nós, aquelas outras pessoas são zumbis, sem vestígio algum de mentalidade. Porque estados mentais não são necessariamente idênticos a estados cerebrais, uma vez que podemos imaginar um sem o outro, eles não são de todo idênticos. Isso significa que a teoria da identidade tem de ser rejeitada. A objeção de Kripke é fatal a ambas as versões da teoria da identidade, tanto a tipo-tipo quanto a espécime-espécime, igualmente. Não pode ser evitada pelo favorecimento da teoria espécime-espécime em detrimento da versão tipo-tipo.<sup>15</sup>

A discussão de Kripke de que é possível conceber ou imaginar uma dor sem a presença de um estado cerebral correspondente (ou, de fato, qualquer estado físico que seja) é uma reminiscência do argumento de Descartes na Sexta Meditação (veja Capítulo 2, seção “O argumento apartir da percepção clara e distinta”) – isto é, de que pode conceber uma situação em que seja possível que sua mente exista sem o seu corpo. Na eventualidade de isso ser o caso, então, como salientei antes, há a possibilidade de que o mental possa existir sem o físico, e isso quer dizer que qualquer tentativa de reduzir o mental ao físico, assim como a temperatura de um gás pode ser reduzida ao movimento molecular, não pode dar bom resultado. A temperatura de um gás não pode existir sem o movimento molecular, porque, em última instância, é dele que consiste a temperatura. Porém, a possibilidade de que estados mentais possam ocorrer sem a presença dos correspondentes estados físicos significa que, em última instância, o mental não pode ser constituído pelo físico. Consequentemente, um materialismo completa e profundamente redutivo, ao final de contas, não pode ser alcançado.

### **Materialismo eliminatório**

Não obstante as objeções contra o materialismo redutivo, que a muitos filósofos parecem ser intransponíveis, não surpreende que ele ainda tenha seus campeões. Nesta seção final, explicarei e comentarei a forma mais radical que o materialismo assume hoje, qual seja, a do materialismo eliminatório. O que distingue esta teoria é que não só pretende mostrar que quando falamos sobre estados mentais estamos realmente falando sobre estados cerebrais, mas que temos de desistir completamen-

te de falar sobre estados mentais. Não só não há estados mentais que existam sobre e acima de itens físicos, mas o próprio conceito de mental e o vocabulário mediante o qual encontra sua expressão têm de ser apagados de nossos hábitos de pensamento e fala. Não há preocupação alguma quanto a como o mental deve ser reduzido ao físico, porque, de acordo com o defensor do materialismo eliminatório, as propriedades mentais não existem, nem nunca existiram. Uma vez falamos de gnomos, bruxas, planeta Vulcano, fluido calorífico, éter, o misterioso "pó da simpatia", que supostamente propiciava a longevidade, e o flogisto. Agora, esses termos não figuram mais seriamente em nosso vocabulário, porque nos demos conta de que eles não se referem a coisa alguma. Nenhum dos itens nomeados por essas palavras existe, ou jamais existiu. É pura mitologia, o resultado de confuso pensamento pré-científico e pré-filosófico.

O mesmo destino aguarda os estados mentais. Não há estados mentais, somente processos cerebrais. Uma vez imaginamos a existência de coisas tais como estados mentais, inclusive consciência, mas estávamos iludidos. Realmente, só o que havia eram processos cerebrais, afinal de contas. Portanto, esse tipo de teoria da identidade é também conhecida como a teoria do desaparecimento. Metade da identidade, a metade mental, por assim dizer, foi absorvida pela metade física (sumiu por passe de mágica, tem-se a tentação de dizer), e o processo foi abolido. Não só os processos mentais segundo sua concepção tradicional não existem, como faremos muito melhor se deixarmos de persistir no hábito de pensar sobre eles como se de fato existissem. Falar sobre estados mentais pertence a uma psicologia popular antiquada (o termo "popular", com suas conotações bucólicas e de simplicidade campesina,

tem a intenção de ser ligeiramente abusivo), e, em vez disso, deveríamos abraçar um vocabulário atualizado, baseado nos termos dos últimos resultados das ciências biológicas, físicas e psicológicas. Termos mentais e os itens aos quais eles alegadamente se referem devem ser evitados, e devemos não mais falar de sentir dor ou ser devastado pelo ciúme, mas, em vez disso, falar de fibras C disparando, ou de circuitos no córtex trino e uno dando descarregas com pontiagudas freqüências de várias ordens de hertz.

Veja o Exercício 3.7 antes de ler minhas próprias posições sobre o materialismo eliminatório.

### **EXERCÍCIO 3.7**

Qual é a sua resposta à proposta do materialismo eliminatório? Quão plausível você acha que essa teoria é? Após haver feito suas próprias reflexões sobre essa questão, combine com alguém para discuti-la. Depois veja a minha resposta.

Pelo teor dos meus comentários sobre o materialismo redutivo ao longo deste capítulo, certamente não será uma surpresa para você saber que acho toda a proposta do materialismo eliminatório inacreditável e ridícula. Acredito que posso estar mais certo de que eu sou consciente e desfruto de uma vida mental rica e variada do que posso estar certo da verdade de qualquer teoria que tente negar inegáveis fatos da minha experiência pessoal. O fato da minha experiência consciente não é uma construção de teoria relacionada a hipóteses equivocadas de cientistas, ou às fantasias de mitologistas, mas constitutivas de minha própria vida como um ser humano. Para negar esses fatos eu teria de fingir estar anestesiado!

Norman Malcolm, criticando Katherine Wilkes, que defende o materialismo

eliminatório em seu livro *Physicalism*<sup>16</sup>, tece um vigoroso comentário em resposta à afirmação que ela faz de que, no futuro, as crianças não mais dirão ter uma dor de cabeça, mas, em vez disso, referirão que suas fibras C estão disparando.

Que solução maravilhosa! Uma expressão é substituída por outra. Isso *eliminaria* o discurso da sensação? De forma alguma. Pois a nova expressão não *tomaria o lugar* da expressão familiar, a menos que a ela fosse atribuído o *mesmo uso*. Por exemplo, a expressão de primeira pessoa "minhas fibras C estão disparando" não teria de ser usada como uma expressão imediata de sensação, não como uma hipótese sobre o que está acontecendo no cérebro do falante, nem como um relato de uma observação feita com o auxílio de instrumentos. Mas, então, nada teria mudado, a não ser um pouco de terminologia.<sup>17</sup>

Em outras palavras, a mera mudança de um pouco de vocabulário não abolirá com isso as várias distinções perceptuais entre os estados mentais e físicos, que exploramos no Capítulo 1.<sup>18</sup>

Poderia ser objetado que isso daria muito pouca atenção a Wilkes. Afinal de contas, ela é uma eminentemente filósofa, que juntamente com outros filósofos, como Paul e Patricia Churchland, têm firmemente advogado a eliminação de termos e categorias mentais em favor dos físicos. Certamente deve haver algo mais se passando no fisicalismo do que o que eu indiquei, se os pensadores que acabei de mencionar estão dispostos a defendê-lo?

Em resposta eu diria que, pelas razões que delineei, posso certamente ver a motivação daquelas que advogam o materialismo eliminatório. Sua adoção não só simplifica a nossa ontologia, mas torna óbvio o problema de como a interação entre mente/corpo é possível, juntamente com

a questão de exatamente como os eventos físicos no cérebro teriam a possibilidade de produzir tal diferente ordem de ser, como o fenômeno da consciência. Essas questões são extremamente difíceis e prementes.

No entanto, a resposta correta a essas questões, creio, é enfrentá-las de frente, e não simplesmente passar ao lado delas, atirando-se aos braços do materialismo eliminatório. Negar que a consciência e outros estados mentais existem é realmente, não obstante ser extremamente improvável que o materialismo eliminatório venha a ser verdadeiro, optar pelo caminho mais fácil. Pois, então, seríamos poupadinhos da difícil tarefa de enquadrar a mente no mundo físico. Ao contrário, a existência da consciência deveria ser tomada com seriedade, e a questão de como a mentalidade deve ser explicada e localizada dentro da nossa visão de mundo não deveria ser obstaculizada. É a esses desafios que me voltarei nos últimos capítulos.

### Questões para reflexão

1. Se meu acesso a meus estados mentais é privilegiado, mas meu acesso a meus estados cerebrais não o é, isso quer dizer que meus estados mentais não podem ser meus estados cerebrais?
2. "Não há objeções lógicas à tese da identidade mente/cérebro." Discuta.
3. Explique a diferença entre a intencionalidade intrínseca e a derivada. Pode um sistema puramente físico possuir intencionalidade?
4. Descreva três traços de estados mentais que, supostamente, os distinguem de estados físicos. Esses traços significam que estados mentais não podem ser inteiramente parte do mundo físico?

\*N. de T.: Em português: *O fisicalismo*.

5. Explique a diferença entre as teorias tipo-tipo e espécime-espécime. Quais vantagens, se houver alguma, a teoria espécime-espécime tem sobre a teoria tipo-tipo?
6. A capacidade de realização múltipla de estados mentais é incompatível com sua redução básica a estados puramente físicos?
7. Que razões existem para que se pense que, embora possamos evitar um dualismo de substâncias, não podemos evitar um dualismo de propriedades?
8. Pode-se atribuir aos pensamentos uma localização literal no cérebro?

### Sugestões para leituras aprofundadas

- Um artigo seminal é o de J.J.C. Smart, *Sensations and the Brain Processes*, originalmente publicado em *Philosophical Review*, LXVIII, 1959, p. 141-56, mas disponível em muitas antologias. Embora mais de 40 anos tenham se passado desde que foi escrito, este artigo ainda merece ser lido, porque expõe as questões com muita clareza.
- BORST, C.V., ed. *The Mind/Brain Identity Theory*. London, Macmillan, 1970 é uma coleção de artigos de grande utilidade, mas que pode estar esgotada e talvez só possa ser obtida em bibliotecas universitárias.
- CHURCHLAND, Paul, *Matter and Consciousness*, Cambridge, MA, e London, MIT Press, 1988, oferece uma breve introdução à teoria da identidade e é útil para que se tenha uma rápida visão panorâmica.
- HEIL, John, *Philosophy of Mind: A Contemporary Introduction*. Lon-
- don, Routledge, 1998 é outra introdução ao materialismo, que não é demasiadamente difícil.
- MacDONALD, Cynthia, *Mind/Brain Identity Theories*, London, Routledge, 1989. Para aqueles que desejam fazer uma pesquisa mais aprofundada; aborda as questões de forma consideravelmente profunda. Os iniciantes podem achar algumas partes meio pesadas.
- KIM, Jaegwon, *Mind in a Physical World*, Cambridge, MA, Cambridge University Press, 1998, apresenta um pano de fundo útil. O primeiro capítulo, "The Mind/Body Problem: Were We are Now" delineia a história recente do pensamento materialista sobre a mente e é muito útil para quem deseja uma rápida apresentação das questões, enquanto o último capítulo aborda o reducionismo em detalhes. Outro livro razoavelmente fácil é KIM, Jaegwon, *Philosophy of Mind*, Boulder, CO, Westview Press, 1996. Este contém boas discussões sobre o materialismo redutivo, e também fornece uma visão geral de várias teorias do corpo e da mente.
- MOSER, P.K. e TROUT, J.D., eds, *Contemporary Materialism*, London, Routledge, 1995 contém uma atualizada coletânea de artigos contemporâneos sobre o materialismo; igualmente, ROBINSON, Howard, ed. *Objections to Physicalism*, Oxford, Clarendon Press, 1993.
- CHALMERS, David, *The Conscious Mind*, Oxford, Oxford University Press, 1996 contém muitas observações e discussões interessantes, mas com 404 páginas (incluindo notas) é quase uma maratona.

## NOTAS

1. SMART, J.C., *Sensations and Brain Processes*. *Philosophical Review*, 68, 1959. Reimpressão: BORST, C.V., ed., *The Mind/Brain Identity Theory*. London, Macmillan, 1970, p. 52-66.
2. NAGEL, T., "Physicalism", in: BORST, C.V., ed. *The Mind/Brain Identity Theory*. London, Macmillan, 1970, p. 218.
3. Compare as colocações de R. Brandt e J. Kim: "A identidade é uma relação simétrica; não favorece um de seus termos às custas do outro". *The Logic of the Identity Theory*, *Journal of Philosophy*, 64, p. 515, 1967.
4. SEARLE, John, *Minds, Brains, and Science. The 1984 Reith Lectures*. London, Penguin, 1989.
5. FODOR, J., *Psychosemantics*. Cambridge, MA, MIT Press, 1987, p. 97.
6. SEARLE, J., *The Rediscovery of Mind*. Cambridge, MA, MIT Press, 1992, p. 51; STRAWSON, G., *Mental Reality*. Cambridge, MA, MIT Press, 1994, c.7; JACQUETTE, D., *Philosophy of Mind*. Englewood Cliffs, NJ, Prentice Hall, 1994, c.4.
7. JACQUETTE, D., *Philosophy of Mind*. p. 96.
8. WITTGENSTEIN, L., *Notebooks 1914-16*. [von WRIGHT G.H. e ANSCOMBE, G.E.M., eds; ANSCOMBE, G.E.M., trad.] Oxford, Blackwell, 1961, p. 82.
9. HACKER, P.M.S. *Insight and Illusion*. Oxford, Clarendon Press, 1972, p. 48, para um relato deste comentário.
10. DAVIDSON, D., in: GUTTENPLAN, S., ed. *The Blackwell Companion to the Philosophy of Mind*. Oxford, Blackwell, 1994, p. 232.
11. Compare: KIM, J., *Philosophy of Mind*. Boulder, CO, Westview Press, 1996, p. 58.
12. Tive muito auxílio na minha compreensão destas questões em ibid. c. 9.
13. SEARLE, J., *The Rediscovery of Mind*. Cambridge, MA, MIT Press, 1992, p. 40.
14. KRIPKE, S., *Naming and Necessity*. Oxford, Blackwell, 1972.
15. Compare: CHALMERS, David, *The Conscious Mind*. Oxford, OUP, 1996, p. 147.
16. Kripke desenvolve o argumento em duas formas diferentes, uma contra as teorias da identidade de espécime e outra contra as teorias da identidade de tipo. As teorias da identidade de espécime sustentam que dores *particulares* (tal como a minha dor agora) são idênticas a estados cerebrais particulares, tais como as fibras C disparando em minha cabeça agora). Kripke argumenta... que uma dor particular poderia ocorrer sem o estado cerebral particular associado, e vice-versa. Consequentemente, eles não podem ser idênticos. As teorias da identidade de tipo sustentam que estados mentais e estados cerebrais são idênticos como tipos: dor, por exemplo, poderia ser idêntica como um tipo ao disparo de fibras C. Kripke mantém que isso é claramente refutado pelo fato de que se poderia instanciar o tipo de estado mental sem o tipo de estado cerebral e vice-versa. Afinal de contas, podemos contar quatro argumentos separados aqui, divididos de acordo com o alvo (teorias de identidade de espécime ou de tipo) e de acordo com o método de argumento (a partir da possibilidade de desincorporação ou da possibilidade de zumbis).
17. ARMSTRONG, D.M. e MALCOLM, N., *Consciousness and Causality*. Oxford, Balckwell, 1984, p.99; ênfase no original.
18. Minha posição é essencialmente a mesma de Karl Popper. Ao escrever sobre o materialismo eliminativo – ao que ele se refere como materialismo promissor – ele diz: "Há racionalidade, não mais interesse a ser encontrado na tese do materialismo promissor do que, digamos, na tese de que um dia aboliremos gatos e elefantes, por cessar de falar a respeito deles; ou na tese de que um dia aboliremos a morte, por cessar de falar nela. (De fato, não nos livraremos das pulgas simplesmente por recursarmos de falar a respeito delas?)." Veja POPPER, K. *The Self and Its Brain*. London, Heidelberg, New York, Springer International, 1977, p. 97.

# 4

## BEHAVIORISMO ANALÍTICO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Compreender a diferença entre behaviorismo psicológico e filosófico ou analítico.
- Compreender a proposta central do behaviorismo analítico, qual seja, a de que falar sobre estados mentais é equivalente quanto ao significado a falar sobre padrões de comportamento reais e potenciais.
- Compreender os pontos fortes do behaviorismo analítico.
- Compreender a versão "dura" do behaviorismo de Hempel e os fatores que a motivaram.
- Compreender e avaliar a importância de evitar a circularidade ao tentar fazer uma redução analítica de estados mentais.
- Compreender as dificuldades com que Hempel se depara na tentativa de fornecer uma análise comportamental não-circular de estados mentais.
- Compreender a versão "branda" do behaviorismo de Ryle e os fatores que a motivaram.
- Compreender a abordagem de Ryle de uma disposição e ser capaz de avaliá-la.
- Compreender a dificuldade com que o behaviorismo se depara ao tentar fornecer explicações de ações e comportamentos em geral.
- Compreender a dificuldade com se depara o behaviorismo para abordar a fenomenologia de estados não-intencionais e sua aparente privacidade.
- Compreender por que o behaviorismo não pode explicar a autoridade de primeira pessoa no que diz respeito ao conhecimento que o indivíduo tem de seus próprios estados mentais.
- Compreender por que a possibilidade de fingimento não pode mostrar que o behaviorismo seja falso, mas avalia a força dos argumentos dos zumbis e superespartanos para atingir esse objetivo.

## Introdução

O dualismo e a teoria da identidade de mente/cérebro compartilham o pressuposto de que a mente seja uma coisa, uma substância não-física cartesianas, por um lado, e um cérebro vivo, por outro. O behaviorismo analítico desafia e rejeita este pressuposto, como o funcionalismo o faz, como será explorado no próximo capítulo. De modo mais direto e simples, o behaviorismo analítico sustenta que afirmações sobre a mente e estados mentais vêm a ser, após a análise, equivalentes a afirmações que descrevem o comportamento público real e potencial de uma pessoa. Há, sob esse ponto de vista, no que concerne aos estados mentais de alguém, em última instância, nada mais do que certos padrões de comportamento observáveis que ele exibe, ou que, em circunstâncias apropriadas, ele se dispõe a manifestar.<sup>1</sup>

O behaviorismo analítico, para que possa ser posto em funcionamento, tem uma quantidade de pontos fortes. Estes são:

**Evitar o problema da interação mente/corpo.** Primeiramente, ele torna perspicua a natureza da relação da mente para com o corpo: a mente é somente o comportamento, real e potencial, do corpo. A mente não causa comportamento: ela é comportamento. Preocupações quanto à natureza da mente, concebida como uma substância imaterial, não podem surgir.

**O caráter não-misterioso do mental.** Não há propriedades mentais misteriosas a serem explicadas. Caso um behaviorista esteja de algum modo preocupado com a relação da mente com o cérebro, o problema se reduz a puramente científico, em que estados físicos internos

são causalmente responsáveis por vários padrões de comportamento. Todas as especulações, quanto a como propriedades não-físicas podem emergir a partir de processos físicos subjacentes, tornam-se redundantes.

**Dissolução do problema de outras mentes.** Apesar de o dualismo cartesiano tornar impossível o conhecimento de estados mentais que não os da própria pessoa. Os estados mentais abarcam um reino radicalmente privado, a que só tem acesso a pessoa à qual pertencem tais estados. O behaviorismo desvia dessa dificuldade, porque ao ver o comportamento de alguém estamos testemunhando a sua mente em ação. Aos animais também se pode atribuir mentes, uma vez que os animais exibem uma gama de comportamentos, desde os simples até os complexos. Desta forma o distanciamento que o cartesianismo colocaria entre nós e o resto do reino animal desaparece.

No entanto, contra seus pontos fortes, há uma quantidade de obstáculos sérios com os quais o behaviorismo analítico deve se deparar, como veremos mais tarde.

## Behaviorismo analítico contrastado com Behaviorismo metodológico

Uma caracterização mais completa do behaviorismo analítico ainda está por ser dada. No entanto, antes que possamos prosseguir, será necessário distinguir o behaviorismo analítico, que é uma posição filosófica, do behaviorismo como ele ocorre na psicologia. Caso essa distinção não seja feita, pode conduzir a certas confusões, especialmente entre aqueles que vieram à filosofia após haver estudado psicologia.

O behaviorismo em psicologia não é uma teoria que diz respeito a como se deve bem compreender o discurso sobre a mente, mas um método de fazer psicologia, uma proposta de como se deve conduzir a pesquisa em psicologia, uma vez que sua completa caracterização é o behaviorismo metodológico. Esta abordagem à psicologia se reflete no fato de que frequentemente a disciplina é definida como a ciência do *comportamento animal e humano*, e não como o estudo de *mentes humanas e de animais*. A ideia é a de que qualquer coisa que tenha a pretensão de se autointitular ciência deve estudar o que é público, porque somente o que é público é passível de ser medido e quantificado e, o que é ainda mais importante, verificável por outros observadores e experimentadores. Os resultados dos experimentos devem ser reduplicados por outros observadores independentes para que tenham alguma chance de ser aceitos como teorias viáveis pelas assim chamadas ciências "duras", como a física, e as mesmas exigências são aplicadas às afirmações e teorias da psicologia.

O behaviorismo, nas mãos dos psicólogos norte-americanos J. B. Watson e B. F. Skinner, assumiu a forma da afirmação de que o comportamento humano deveria ser compreendido como um conjunto de respostas evocadas por estímulos externos. Os processos "internos", na forma de eventos físicos cerebrais ou de eventos mentais, eram postos de lado, e se afirmava que o conhecimento dos estímulos externos e o comportamento que estes causavam eram tudo de que se necessitava para explicar por que as pessoas se comportavam do modo como faziam. Skinner, em particular, acreditava que o uso de termos intencionais em psicologia engajava a pessoa na crença da existência de um homúnculo, um pequeno homem na cabeça, e isso conduz à conhecida ob-

jeção de que recursos a homúnculos não explicam coisa alguma, uma vez que as habilidades e o comportamento do próprio homúnculo demandam explicações (veja Capítulo 2, seção "A falácia do homúnculo"). Então, em vez de recorrer às crenças, às intenções e aos desejos de uma pessoa para explicar seu comportamento, vamos explicá-lo em termos de respostas condicionadas que foram reforçadas por repetição e recompensa. O comportamento de um pombo pode ser "formatado" pelo estímulo de reforço de uma recompensa em alimento a cada vez que fizer um tipo particular de movimento; assim, finalmente, por esse meio, ele pode ser treinado a caminhar em círculos. Do mesmo modo, toda a conduta humana supostamente pode ser explicada puramente em termos de relações de estímulo-resposta reforçadas, produzidas por tal condicionamento em operação.

Tendo concluída esta breve apresentação do behaviorismo metodológico, devemos agora voltar a analisar mais atentamente a variedade filosófica, que é o nosso verdadeiro objetivo.

### Behaviorismo analítico

A análise de todo o discurso sobre a mente em termos descritivos do comportamento real e possível é uma forma de reducionismo, mas representa uma variedade diferente daquele tipo que encontramos anteriormente em relação à teoria mente/cérebro. Os teóricos da identidade mente/cérebro nunca desejaram sustentar que o discurso sobre estados mentais era equivalente, quanto a seu significado, ao discurso sobre estados cerebrais. Conceitos mentais eram irredutivelmente diferentes quanto ao significado de conceitos físicos, dos quais o discurso sobre o cérebro e seus estados forma uma sub-

classe. No entanto, no nível dos itens ou estados de coisas que se enquadram nos conceitos, a teoria da identidade mente/cérebro alega que não há quaisquer fatos mentais separados e além de fatos de estados cerebrais, estes sendo constituídos por fatos físicos envolvidos nas operações materiais do cérebro. No caso da teoria da identidade, a situação é similar àquela da relação entre a temperatura de um gás e a energia cinética das moléculas que o compõem. O discurso sobre a temperatura e aquele sobre a energia cinética das moléculas não são sinônimos: foi necessária pesquisa empírica para se descobrir a identidade entre esses dois fenômenos aparentemente diferentes, e não meramente a reflexão sobre os significados das palavras e sua respectiva análise. Ainda assim, não há fatos sobre a temperatura separados e além dos fatos sobre a energia cinética de moléculas.

O behaviorista analítico, por outro lado, insiste que os enunciados que descrevem estados mentais ou psicológicos podem ser traduzidos, sem perda de significado, para enunciados que descrevem comportamento real e possível. A ideia é que tudo que pode ser dito no vocabulário familiar de sensações, crenças, pensamentos, consciência, e assim por diante, pode ser de alguma forma melhor dito, ou dito de maneira mais perspicua, em outros termos, não-mentais, a saber, as descrições comportamentais. Essas descrições comportamentais devem ser pensadas como reveladoras da verdadeira natureza do mental. O behaviorismo analítico assemelha-se a outras formas de reducionismo, como, por exemplo, o fenomenalismo, que sustenta que tudo que dizemos sobre objetos físicos pode ser expresso em termos de experiências reais e possíveis que um observador ideal dos objetos em questão apreciaria. O discurso sobre mesas, para tomar uma ilustração banal da tese, deve

ser traduzido em enunciados descrevendo experiências reais e possíveis de mesa, que um observador ideal, adequadamente posicionado, teria. Similarmente, o reducionismo acerca de enunciados sobre sonhos nega que sonhos sejam experiências privadas que ocorrem quando as pessoas estão dormindo à noite, e são mais bem entendidos como enunciados que descrevem suas experiências matinais, quando estão despertas.<sup>2</sup>

É importante perceber, logo de saída, que se esses padrões de análise devem ter qualquer chance de sucesso, então os termos que constituem a análise não devem conter, ou pressupor, nada do vocabulário mental que está sendo analisado, pois, de outro modo, a análise será circular. Com efeito, ela estará apenas se repetindo.

Para o caso de esse ponto demasiadamente abstrato mostrar-se difícil de compreender à primeira vista, eu ofereço uma comparação, no Exercício 4.1, que você deve analisar antes de prosseguir.

#### **EXERCÍCIO 4.1**

Eu quero fornecer uma análise do conceito de uma causa para explicar o que significa falar em causa e efeito, e, consequentemente, proponho a seguinte fórmula: "A é a causa de B" deve ser analisado como "B acontece por causa de A". Você é capaz de ver por que essa tentativa de análise não funcionará?

Eu penso que é bastante claro por que a análise fornecida no Exercício 4.1 deve ser rejeitada, uma vez que o segundo enunciado que constitui a alegada análise é meramente uma paráfrase do primeiro. Isto é, ele emprega o próprio conceito de causa, que supostamente deveria estar sendo analisado. Dizer que "B acontece por causa de A" é meramente uma outra maneira de dizer "A causa B". A situação

não melhoraria se, em vez da sugestão inicial, eu propusesse o enunciado "B é o efeito ou resultado de A", ou mesmo "A faz B ocorrer". Todos esses enunciados, eu sugiro, não nos levam para além do que é dito no enunciado "A causa B", como o qual começamos.

Uma análise famosa da causalidade, apresentada por David Hume (1711-1776), leva em consideração a necessidade de que a análise de um conceito ou fenômeno seja genuinamente formulada em outros termos, para que a análise seja informativa e pretenda produzir uma compreensão melhor e mais clara. A primeira tentativa de Hume em relação a "A causa B" foi "Sempre que A ocorre, segue-se B". A causalidade, para usar a expressão do próprio Hume, é a conjunção constante de certos tipos de eventos, uma regularidade que descobrimos com base na experiência repetida de um tipo de evento, em um dado conjunto de circunstâncias, sendo invariavelmente seguido por um outro tipo. Não podemos deduzir logicamente um efeito de sua causa, ou uma causa de seu efeito. Meramente descobrimos, como uma questão de fato contingente, que os dois tipos de eventos que constituem a relação causal se relacionam desse modo. Isso tem a importante consequência adicional de que, se dois eventos estão causalmente relacionados, deve ser possível identificar cada um dos eventos de maneira completamente independente da identificação do outro. Em outras palavras, deve haver descrições dos eventos que não carreguem consigo implicações lógicas para a existência ou natureza de nenhum outro evento. Deveremos retornar a esse importante ponto posteriormente, no Capítulo 6, seção "Searle e o problema mente/corpo", assim como no Capítulo 7, seções "Razões para ações como causas" e "A resposta de Davidson: a teoria causal da ação intencional".

Embora essa primeira análise da causalidade não seja adequada, dado que é possível pensar em muitos exemplos nos quais um tipo de evento é regularmente seguido de outro tipo, mas um não causa o outro, ela é ao menos genuína e informativa e não comete o erro fundamental de meramente repetir-se.<sup>3</sup>

Uma maneira abreviada para rotular convenientemente esse erro, usando algumas expressões latinas, é a seguinte. O fenômeno a ser analisado é chamado de *analysandum*, e a análise do fenômeno é chamada *analysans*. Mas, sob pena de circularidade (e, em última instância, de trivialidade), o *analysans* não deve conter o *analysandum*. Um erro paralelo pode ocorrer com as definições: o *definiens* (a definição de um termo) não deve conter o *definiendum* (o termo a ser definido). O programa do fenomenalismo, mencionado anteriormente, foi por fim julgado incapaz de ser bem-sucedido por parte dos filósofos porque se descobriu que, mais cedo ou mais tarde, o *analysans* teria de incorporar a própria noção de um objeto físico, o *analysandum*, que supostamente deveria estar sendo expresso em termos puramente experenciais.

Entretanto, para voltar à discussão principal, o desafio enfrentado pelo candidato a behaviorista analítico é produzir uma análise de termos mentais usando um vocabulário não-mentalista que empregue apenas descrições comportamentais, sem conter expressão alguma que englobe, ou pressuponha, quaisquer conceitos psicológicos. Mas o que é exatamente uma descrição comportamental?

### Behaviorismo "duro" de Hempel

Uma resposta a essa questão, assim como uma compreensão de uma possível

motivação para se adotar o behaviorismo, é fornecida pela contribuição de Carl Hempel ao debate. Hempel (1905-1997) foi um membro do Círculo de Viena, os positivistas lógicos, que estavam ativos principalmente nas décadas de 1920 e 1930. Os positivistas lógicos estavam bastante interessados na natureza do conhecimento e da metodologia científicos e tinham como uma de suas principais metas a obtenção de alguma espécie de unidade científica. Para eles, todo conhecimento e explicação genuínos deveriam, em última instância, ser compreendidos de acordo com o modelo das ciências físicas. A psicologia deveria ser unificada com a ciência "dura" da física por meio de redução a esta última. Essa concepção do conhecimento era informada pela adoção do princípio de verificação, sendo dela consequência. De acordo com esse princípio, a menos que um enunciado pudesse ser verificado empiricamente (deixando-se de lado as verdades analíticas, não-empíricas, da lógica e da matemática), ele deveria ser rejeitado como sendo desprovido de significado, como estando literalmente vazio de qualquer significância. O enunciado em questão não precisava ser verificado na prática, mas apenas em princípio. Por exemplo, nós não podemos até o momento descobrir se é verdade que há planetas na grande nebulosa de Andrômeda porque ela está muito distante. Mas nós sabemos que tipos de coisas teríamos que fazer para verificar se há planetas por lá – construir um telescópio mais poderoso, enviar uma sonda espacial, e assim por diante.

A teoria verificacionista do significado tinha a seguinte consequência. Se interpretarmos os enunciados sobre estados mentais como dizendo respeito a estados de coisas logicamente privados, acessíveis em princípio apenas pelo possuidor daqueles estados, como sustenta o dua-

lismo, então os enunciados sobre estados mentais de outros não podem, mesmo em princípio, ser verificados por terceiras pessoas. Mas então o discurso sobre outras mentes se tornaria sem sentido. De fato, embora se possa alegar que alguém pode verificar a existência e a natureza de seus próprios estados mentais, os positivistas provavelmente desqualificariam essa verificação como sendo genuína, com base no fato de que ela não está aberta a um controle público, intersubjetivo. Dado, contudo, que o discurso sobre o estado mental dos outros, assim como o nosso próprio, manifestamente é significativo, a única maneira em que isso pode ser obtido é se estados mentais forem entendidos como consistindo no que é indisputavelmente verificável, a saber, o comportamento externo, público.

É importante, contudo, compreender o que se quer dizer com o termo "comportamento". Examine agora o Exercício 4.2, antes de prosseguir.

#### **EXERCÍCIO 4.2**

Considere, em primeiro lugar, os enunciados (a) "Martin ergueu seu braço" e (b) "O braço de Martin ergueu-se". Eles não são equivalentes no que toca ao significado. Você pode notar a diferença de significado entre esses dois enunciados?

A razão para a não-equivalência dos enunciados (a) e (b) no Exercício 4.2 é a seguinte. O enunciado (a) relata algo que Martin fez, uma ação que ele desempenhou, presumivelmente de modo intencional, e comumente com algum propósito em mente. O enunciado (b) não relata algo que Martin fez; ele meramente descreve o movimento do braço de Martin pelo espaço, deixando em aberto, na melhor das hipóteses, a questão sobre se

o erguer do braço foi efetuado por Martin ou não, e, na pior das hipóteses, implicando que o movimento do braço foi *meramente* um movimento, isto é, que não foi algo que Martin fez. O enunciado (a) implica o enunciado (b), mas (b) não implica (a). Se Martin ergueu seu braço, então se segue que seu braço ergueu-se. Mas se o braço de Martin ergueu-se, não se segue que ele o ergueu. Seu braço poderia ter sido movido por razões completamente desconectadas de sua capacidade de agir. Por exemplo, uma razão externa para seu braço mover-se poderia ser o fato de que alguém o ergueu. Igualmente, pode ter havido uma razão interna, um espasmo muscular, talvez, ou o disparo involuntário das fibras nervosas que controlam o movimento.

Isso quer dizer que há dois modos distintos de descrições comportamentais possíveis. O primeiro é constituído por descrições do que as pessoas fazem, das ações que elas desempenham, e dos feitos que elas realizam por meio de sua capacidade pessoal de agir. Podemos chamá-las, para facilitar a referência, de descrições agenciais. O segundo modo de descrição caracteriza o que acontece em termos de movimentos corporais, frequentemente por implicação e contexto, excluindo-os da esfera da ação, mas, por outras vezes, deixando em aberto a questão sobre se esses movimentos realmente incluem ações, ou se eles devem ser vistos como meros movimentos corporais – movimentos corporais “incolors”, para empregar a útil expressão do psicólogo C. L. Hull. Descrições de movimentos corporais “incolors” poderiam, em teoria, ser refinados até virarem descrições sofisticadas da matéria em movimento, empregando-se termos e conceitos de matemática, geometria e física.

Assim, em vez de dizer, por exemplo, que Suzanne apertou sua bochecha, poderíamos dizer que se observou que certa

porção de matéria de tais e tais dimensões (sua mão e seu braço) foi movida de um conjunto de coordenadas no espaço para outro conjunto de coordenadas durante um período de tempo especificado. Similarmente, em vez de dizer que alguém sortiu, teríamos de dizer que um pedaço de carne de certas dimensões sofreu mudanças em sua forma, caracterizando essas alterações na linguagem da matemática e da topologia.<sup>4</sup>

Você deve agora examinar o Exercício 4.3 antes de prosseguir a leitura.

### EXERCÍCIO 4.3

Você pode perceber que modo de descrição Hempel deve escolher para sua análise comportamental e por quê? Dê razões para sua decisão.

Sob pena de trazer circularidade à análise que ele propõe, Hempel deve abster-se do uso de descrições de movimentos corporais incolors. Isso também estaria de acordo com sua declarada meta de reduzir a psicologia à física. Nessa mesma linha, Hempel escreveu: “Nós vemos claramente que o significado de enunciados psicológicos consiste apenas na função de abreviar a descrição de certos modos de resposta física, característicos dos corpos de homens e animais”.<sup>5</sup> Deve-se dizer, contudo, que a análise comportamental tentada por Hempel, quando apareceu, de uma afirmação psicológica, “Paulo tem uma dor de dente”, ficou aquém de ser bem-sucedida, violando instantaneamente a restrição segundo a qual nenhuma análise deveria conter os próprios termos que ela supostamente deveria analisar. A análise proposta por Hempel compunha-se de cinco elementos:<sup>6</sup>

1. Paul chora e faz gestos de tais e tais tipos.

2. Confrontado com a questão "O que há de errado?", Paul profere as palavras "Eu tenho uma dor de dente".
3. Um exame mais detido revela um dente cariado com a polpa exposta.
4. A pressão sanguínea de Paul, seus processos digestivos, a velocidade de suas reações, indicam tais e tais mudanças.
5. Tais e tais processos ocorrem no sistema nervoso central de Paul.

As condições (3), (4) e (5) não parecem essenciais para uma análise do significado do enunciado "Paul tem uma dor de dente". Poderia ser verdade que Paul tivesse uma dor de dente, mesmo que um exame físico não revelasse nenhum dente cariado e nenhuma mudança em sua pressão sanguínea, em seus processos digestivos ou no seu sistema nervoso. Inversamente, as mudanças físicas recém-descritas poderiam ocorrer, e ainda assim Paul poderia não sentir dor. É claro que isso pode ser enigmático, mas não seria autocontraditório conjugar o relato verídico de Paul de que ele tinha dor de dente com a negação de que qualquer coisa fora do comum estivesse acontecendo com seu dente ou com seu corpo. (Afinal, pense nas pessoas que possuem as assim chamadas dores psicosomáticas, para as quais nenhuma causa física óbvia está disponível.) Isso é suficiente para mostrar que (3), (4) e (5) não podem ser parte do que significa Paul ter uma dor de dente.

Isso nos deixa com (1) e (2). Em relação a (1), o problema é que termos como "chora" e "gesticula" já implicam atribuir estados mentais a Paul. "Chora" implica que Paul está sofrendo alguma experiência desagradável (uma descrição mais neutra seria "está saindo água dos olhos de Paul"), e "gesticula" sugere movimentos das mãos feitos por Paul com a intenção de indicar o lugar e a inten-

sidade da dor, e talvez de obter ajuda e simpatia.

Em relação a (2), uma dificuldade adicional que contagia a análise é que é claro que Paul não pode responder à questão a menos que a compreenda. Além disso, ele deve afirmar as palavras que saem de sua boca, isto é, ele deve saber o que as palavras significam e ter a intenção de, com elas, responder à questão do falante.<sup>7</sup> Mas compreender, saber, afirmar e ter uma intenção são todos termos não analisados comportamentalmente. Por fim, Paul só responderá à questão do modo como o faz se ele quiser falar a verdade. Pode ser que Paul esteja sofrendo, mas não de dor de dente, e queira esconder a razão real de sua aflição. Mas "querer" é novamente um termo mental ou psicológico, do qual a análise comportamental deveria abster-se, com o fim de evitar a acusação de circularidade.

### Especificando padrões de comportamento

Uma tentativa séria de levar a cabo a análise de termos mentais de uma maneira não-circular tem de enfrentar numerosas dificuldades. O primeiro problema diz respeito a que padrões de comportamento, caracterizados puramente como movimentos corporais, como o movimento de matéria pelo espaço, devem ser incluídos na análise. A raiz do problema está no fato de que não há correspondência clara, um a um, entre os tipos de ações e os tipos de movimentos corporais. Isso lembra a dificuldade enfrentada pelos teóricos da identidade tipo-tipo mente/cérebro, um dado tipo de estado mental poderia ser idêntico ao disparo de fibras C no cérebro de Keith, mas o disparo de fibras Z no cérebro de Zielfa, ou mesmo o disparo de fibras P no cérebro de Pippa.

Embora todo espécime de um tipo de estado mental seja idêntico a um espécime de algum tipo de estado cerebral, não se trata necessariamente de um e o mesmo tipo de estado cerebral em todas as ocasiões: um espécime de um dado tipo de estado mental é multiplamente realizável em diferentes tipos de estados cerebrais. O mesmo vale para ações. Todo espécime de um tipo de ação será idêntico a um espécime de algum tipo de movimento corporal (movimento inclui imobilidades), mas não é necessário que cada espécime de um dado tipo de ação deva ser corporificado exatamente no mesmo tipo de movimento corporal em todas as ocasiões em que ele ocorre.

Antes de prosseguir a leitura, examine a questão no Exercício 4.4.

#### **EXERCÍCIO 4.4**

Você pode perceber a ameaça imposta pela múltipla capacidade de realização de ações à análise comportamental dos estados mentais?

A ameaça sugerida no Exercício 4.4 é que a análise nunca poderia ser levada a cabo completamente. Para ver isso, considere novamente os primeiros dois itens na análise comportamental que Hempel tinha em vista para a frase "Paul tem uma dor de dente". Hempel expressa isso como "Paul chora, faz vários tipos de gestos e, quando confrontado com a questão sobre o que há de errado, proferre as palavras 'Eu tenho dor de dente'". Bem, Paul pode fazer essas coisas, mas há uma variedade indefinida de outras coisas que ele poderia fazer em vez delas, e que seriam igualmente consideradas expressões comportamentais da dor que ele sente. Ele poderia, por exemplo, gritar, gemer, procurar a garrafa de uísque, aplicar óleo de cravo-da-índia, sentar-se

e permanecer em um grave silêncio, ou socar sua cama. Se ele for um francês monoglotá, ele não proferirá as palavras em português "Eu tenho dor de dente", mas o equivalente em francês. Por outro lado, se ele for um filipino, ele poderá simplesmente emitir o lamento "Agoy!". Lembre-se: se a análise for bem-sucedida, ela deve mencionar todos aqueles tipos de comportamento, e apenas aqueles tipos de comportamento que são capazes de constituir a expressão de dor de Paul. Isso sugere que, seguindo-se a proposta de Hempel, a análise terá mais ou menos a seguinte apresentação:

Paul tem uma dor de dente = Ou Paul está gemendo ou se contorcendo ou lágrimas estão saindo de seus olhos ou está proferindo as palavras "Eu estou com dor" ou seus vários equivalentes estrangeiros ou ele está procurando a garrafa de uísque ou...

Há pelo menos duas coisas erradas com essa análise. Em primeiro lugar, parece que ela nunca poderia ser completa: há uma variedade indefinida de coisas que Paul poderia fazer. Em segundo lugar, os itens nessa lista mencionam ações. Mas descrições de ações são termos que implicitamente se referem a estados mentais e estão proibidos de serem incluídos na análise. Isso significa que cada tipo de descrição de ação, por sua vez, deverá ser formulado por meio de uma disjunção longa e complexa dos vários tipos de movimentos corporais, por meio dos quais a ação poderia, teoricamente, manifestar-se. Assim, há uma forte aparência de que a análise nunca poderia ser levada até uma conclusão final.

Considere também que, se parece difícil o problema de especificar quais comportamentos devem constituir a análise do simples relato de uma sensação tal como uma dor de dente, imagine o quão

mais difícil seria ir adiante com tal análise quando ela deve referir-se às manifestações comportamentais de uma crença complexa e abstrata. Suponha que eu acredite que ninguém até o momento foi capaz de explicar o que é a consciência e por que ela existe. Como essa crença deveria ser analisada em termos comportamentais? Uma maneira óbvia seria, talvez, por meio da minha expressão verbal da crença por meio do proferimento: "Eu acredito que ninguém até o momento foi capaz de explicar o que é a consciência e por que ela existe". Mas já notamos que tal análise veladamente contém itens mentais e, portanto, deve ser rejeitada. No entanto, mais importante é que deveria estar clara a existência de uma variedade massivamente indefinida de modos por meio dos quais eu, ou alguma outra pessoa, poderia expressar a crença sobre a intratabilidade do fornecimento de uma explicação da consciência, e isso significa que a perspectiva de se obter uma análise bem-sucedida recua ainda mais.

Contudo, de acordo com John Foster, essa não é necessariamente uma objeção conclusiva. Ela pode apenas mostrar que "no caso de pelo menos alguns (e talvez de todos) os enunciados psicológicos, os itens linguísticos que tornam explícitos seus conteúdos em última instância não-mentalísticos são de um tamanho infinito, o que mostra que o próprio conteúdo, em sua forma completamente analisada, possui uma complexidade infinita".<sup>8</sup>

Tanto quanto posso entender desse assunto, o que Foster está defendendo é que isso não quer dizer que a noção de uma análise não-mentalista dos itens mentais seja incoerente, apenas porque, por sua própria natureza, ela jamais poderia ser abarcada por nós. Tudo que isso talvez mostre, afinal de contas, é que há coisas que podemos dizer por meio de um vocabulário mentalista que não somos

capazes de dizer puramente por meio do uso de uma presumida tradução fisicalista daquele vocabulário.

Outro obstáculo à análise que devemos considerar é este: como devemos identificar quais tipos de movimentos corporais incolores devem ser incluídos nela, e quais devem ser excluídos? Não há nada na descrição dos movimentos corporais incolores como tais que forneça uma pista para sua possível inclusão ou exclusão em qualquer padrão particular de análise comportamental de um item mental. Em outras palavras, considerados em si mesmos, do ponto de vista da ciência física, os vários tipos de movimentos corporais não formam nenhum agrupamento natural ou óbvio, de tal modo que pudéssemos perceber neles quais deveriam ser incluídos e quais deveriam ser excluídos de uma tradução comportamental de um estado mental particular. Aparentemente, sua identificação poderia ser abordada somente pela identificação do conjunto possível de ações selecionadas para constituir o estágio inicial da análise. A ideia é esta: pense no conjunto de ações que mais plausivelmente poderia servir como manifestação comportamental de um tipo particular de estado mental. Então, refletindo sobre essas ações, tente imaginar os movimentos corporais brutos que deveriam figurar em sua execução, e aqueles que seriam excluídos. Por exemplo, considere a ação "fazer um gol chutando a bola na rede". Vários tipos de movimentos da perna poderiam figurar na análise, mas não, digamos, os movimentos da cabeça. Contudo, o resultado para o qual estamos nos dirigindo é que a identificação dos movimentos deve proceder via identificação da ação. É a ação que fornece ao grupo de movimentos sua unidade, e é apenas pela via da ação que o agrupamento é identificável e inteligível. Mas isso quer dizer que a análise depende de uma referência

indispensável a um item mentalístico para sua execução, e todo tipo de referência a tais itens foi precisamente o que havia sido excluído da análise sob pena de circularidade.<sup>9</sup>

### Circularidade e regresso infinito

As ações que desempenhamos não resultam de estados mentais singulares, mas de combinações de estados mentais. Por exemplo, suponha que Manjit acredite que a loja do outro lado da rua possua cigarros, mas ele acabou de largar o vício do fumo. Claramente, ele não atravessará a rua com a intenção de entrar na loja e comprar cigarros. Por outro lado, se ele não tivesse largado o vício do fumo e quisesse fumar, então isso, combinado com a crença de que cigarros podem ser encontrados na loja, o levaria a atravessar a rua.

Com isso em mente, vamos supor que queiramos analisar comportamentalmente o enunciado mentalístico "Manjit quer uma carteira de cigarros". Como essa análise procederá? Podemos tentar formular isso dizendo que Manjit irá e comprará alguns cigarros. Mas claramente isso é simples demais. Ele não fará isso, se sua intenção de largar o vício do fumo for mais forte do que seu desejo de fumar, ou se ele acredita que não tem dinheiro suficiente para pagar os cigarros, ou que a loja está temporariamente fechada ou sem cigarros, e assim por diante. Assim, pareceria que qualquer tentativa de explicitar o desejo de Manjit por cigarros, em termos puramente comportamentais, envolverá uma referência a outros estados mentais, nesse caso, um certo número de crenças. Portanto, obviamente, não podemos permitir que a análise contenha termos mentais não analisados, tais como "crença", e por

isso algum modo de detalhar o conteúdo das crenças em termos de comportamento deverá ser encontrado. O problema é que parece que qualquer tentativa de explicitar como Manjit provavelmente se comportará, dado que ele acredita que a loja vende cigarros, deve envolver uma referência aos desejos de Manjit. Sua crença isoladamente não pode motivá-lo à ação, isto é, a comportar-se de uma certa maneira. Para que Manjit atravesse a rua e compre cigarros, a crença precisa ser combinada não apenas com outras crenças, mas também com um desejo por cigarros. De fato, o assunto é ainda mais complicado, porque o desejo de Manjit de comprar cigarros somente levará à ação caso ele seja mais forte do que seus outros desejos, como o desejo de não ser atropelado ao cruzar uma rua movimentada, ou o desejo de evitar cruzar a rua sob a escaldante luz do sol. Assim, parece que uma análise comportamental de um desejo deve envolver referência não apenas a outros desejos, mas particularmente a crenças vistas como termos mentais não analisados. E, do mesmo modo, uma análise comportamental de uma crença deve envolver uma referência não apenas a outras crenças, mas também aos desejos tomados como itens mentais não analisados.

Há duas circularidades envolvidas aqui, uma maior e outra menor. A maior circularidade reside no fato de que não se pode permitir que análise comportamental alguma contenha termos mentais não analisados, e o problema é que, não importando o quanto longe possamos ir na tentativa de converter termos mentais em termos de descrições comportamentais, sempre restará um resíduo de itens mentais não analisados, demandando ainda mais análise comportamental. Parece que esse processo não terá fim, embora se possa argumentar que, em última instância,

isso não comprometa a análise, se Foster estiver correto.<sup>10</sup>

A menor circularidade consiste no fato de que uma análise comportamental de um desejo deve fazer uso da noção não analisada de uma crença, e a análise comportamental de uma crença deve fazer uso da noção não analisada de um desejo. O desejo de Manjit por um cigarro consiste em, dentre outras coisas, ele atravessar a rua para chegar à loja, dado que ele acredita que lá se vendem cigarros; e sua crença de que na loja se vendem cigarros consiste em, dentre outras coisas, ele atravessar a rua para chegar à loja, dado que ele deseja um cigarro. Em outras palavras, a explicação do que é para Manjit querer um cigarro envolve a referência às crenças de Manjit, e a explicação do que é para Manjit possuir uma crença sobre a disponibilidade de cigarros envolve uma referência aos desejos de Manjit. Cada metade da análise inclui, por assim dizer, o resíduo da outra metade.

No entanto, pode-se argumentar que essa circularidade menor possa ser evitada. Sydney Shoemaker argumenta que o problema é análogo ao caso de alguém tentar fornecer uma explicação do que são cargas elétricas positivas:

Imagine alguém cujas crenças sobre cargas elétricas sejam resumidas na seguinte frase: "Coisas com carga positiva atraem coisas com carga negativa e repelem coisas com carga positiva; coisas com carga negativa atraem coisas com carga positiva e repelem coisas com carga negativa; e uma carga negativa pode ser induzida esfregando-se um pente de borracha na lã". E suponha que ele seja instado a definir "carga positiva" e "carga negativa" em termos desse conjunto de crenças. Pode parecer que não haveria como ele conseguir isso sem incorrer em circularidade, uma vez que cada um dos tipos de corrente é caracterizado em termos de sua relação para com o outro.<sup>11</sup>

A solução, afirma Shoemaker, segundo uma proposta de David Lewis, é fazer uso de uma noção de uma "frase de Ramsey", assim chamada porque foi inventada pelo filósofo de Cambridge Frank Ramsey (1903-1930). (Deveremos nos deparar com o uso de uma frase de Ramsey novamente quando considerarmos o funcionalismo no Capítulo 5, na seção "Funcionalismo e reducionismo", em que será empregada com um propósito semelhante.)

Uma frase de Ramsey é construída assim. Os termos "carga positiva" e "carga negativa" são eliminados ao serem substituídos pelos termos "propriedades F e G". A tentativa original de definir as cargas positiva e negativa iva desse modo transforma-se em: "Para algumas propriedades F e G, coisas com F atraem aquelas com G e repelem aquelas com F; coisas com G atraem coisas com F e repelem coisas com G; e G pode ser induzida esfregando-se um pente de borracha na lã".

A expressão "Para algumas propriedades F e G", que precede essa definição, é outra maneira de expressar o que os lógicos chamam de quantificadores existenciais  $\exists(F)$  e  $\exists(G)$ , que devem ser lidos, respectivamente, como "Há algumas coisas F tal que..." e "Há algumas coisas G tal que...". F e G não mencionam as propriedades pelo nome, e, por conseguinte, a circularidade pode ser evitada, uma vez que os termos sendo definidos não ocorrem no *definiens* de cada definição (veja a seção "Behaviorismo analítico" caso você tenha esquecido o significado desse termo).

### O behaviorismo "brando" de Ryle

Nesta seção quero dizer algo sobre a importância das disposições no behavio-

rismo analítico, e abordarei esse tópico por meio de uma breve exploração do trabalho de Gilbert Ryle (1900-1976), a quem a noção de uma disposição é particularmente associada. O behaviorismo de Ryle, na medida em que pode ser classificado como behaviorismo, encontra-se em seu livro *The Concept of Mind\**, publicado originalmente em 1949.<sup>12</sup>

A motivação para a explicação de Ryle acerca da natureza da mente é bastante diferente da de Hempel. Em primeiro lugar, não há evidência alguma no livro de Ryle de que ele tenha subscrito à teoria verificacionista do significado delineada anteriormente. Isso é auspicioso, dado que o trabalho filosófico subsequente sobre esta questão mostrou que o verificacionismo é uma explicação dúbia do conteúdo semântico, confundindo fatalmente a capacidade de especificar que estado de coisas, caso tenha ocorrido, tornaria um enunciado factual verdadeiro, com a exigência bastante diferente de se ser capaz de distinguir se o estado de coisas em questão aconteceu ou não. Apenas a primeira dessas duas exigências pode ser vista razoavelmente como explicitando o que é essencial para que um enunciado factual tenha significância.

Em segundo lugar, Ryle não era um fisicalista buscando reduzir a psicologia à física. Ele suspeitava tanto das teorias fisicalistas da mente quanto desdenhava do dualismo cartesiano, buscando dissolver, em vez de resolver, o problema mente/corpo, abordando a questão de uma nova maneira: "O consagrado contraste entre mente e matéria será dissipado, mas não será dissipado por nenhuma das absorções igualmente consagradas da Mente pela Matéria ou da Matéria pela Mente, e sim de uma maneira muito diferente."<sup>13</sup>

Na introdução ao seu livro, Ryle enuncia que sua professada meta é determinar as coordenadas lógicas cruzadas dos conceitos do mental, de modo a capacitar as pessoas que já podem falar com sentido *com* esses conceitos a poderem, adicionalmente, falar com sentido *sobre* eles, à maneira de um cartógrafo lógico ou filosófico que busca obter uma visão sinótica dos conceitos, tornando claras as suas inter-relações e as regulações que governam seus usos.

O que isso significa, na prática, é que a meta de Ryle é demolir a concepção cartesiana da mente como uma entidade fantasmagórica, não-física, que existe além do ser humano vivo familiar, de carne e osso, uma entidade cujos estados supostamente seriam logicamente privados, no sentido explicado anteriormente (Capítulo 1), aos indivíduos que possuem a mente em questão. Contudo, seria perfeitamente possível repudiar o dualismo, sem por isso ser forçado a abandonar a crença na privacidade radical dos estados mentais, mesmo aceitando-se que um compromisso com o dualismo implica um compromisso com a aceitação da privacidade do mental. Consequentemente, devo considerar o ataque de Ryle a esses aspectos gêmeos do dualismo separadamente.

A primeira meta de Ryle é repudiar completamente o conceito cartesiano da mente como uma substância imaterial, conectada durante a vida a uma máquina corpórea, o corpo. Ryle associa isso a, como ele mesmo diz, "abuso deliberado", como "O Dogma do Fantasma na Máquina." Essa concepção, alega Ryle,

é inteiramente falsa, e falsa não apenas no detalhe, mas em princípio. Não se trata meramente de um agrupamento de erros particulares. Trata-se de um grande

\*N. de T.: Em português: *O conceito de mente*.

erro, e de um erro de uma espécie particular. Trata-se de um erro categorial. Ela representa os fatos da vida mental como se eles pertencessem a um tipo ou categoria lógica (ou a uma gama de tipos ou categorias), quando eles na verdade pertencem a uma outra.<sup>14</sup>

O erro em que Ryle está pensando é este. Não há, em adição aos seres humanos vivos, cujo comportamento externo e mecanismo interno são tão públicos e abertos ao escrutínio quanto as carteiras de árvores, cristais e planetas, entidades fantasmagóricas chamadas mentes, cujos mecanismos internos, "paramecânicos", são acessíveis apenas às pessoas a quem essas mentes pertencem. Uma pessoa não vive duas histórias paralelas, uma consistindo em atos físicos externos, públicos, e a outra consistindo em acontecimentos fantasmagóricos em um palco mental privado.

Para ilustrar o argumento que está apresentando, Ryle propõe uma analogia que se tornou famosa. Devemos imaginar um visitante estrangeiro a quem é mostrada a Universidade de Oxford. Ele visitá os vários colleges, a Biblioteca Bodleiana, os laboratórios, as Escolas de Exame na High Street, a biblioteca da subfaculdade de Filosofia, e assim por diante. No final do seu passeio, ele afirma: "Muito obrigado. Mas, infelizmente, ainda não me mostraram a própria universidade". Seu erro é óbvio. Ele está pensando na universidade como se ela fosse um prédio, assim como os colleges individuais ou a escola de exames, e como se ela fosse capaz de existir independentemente desses, em vez de perceber que a totalidade do que ele viu é a universidade.

O mesmo erro é evidente, de acordo com Ryle, na suposição de que, em acréscimo aos desempenhos observáveis, públicos, dos seres humanos, às ações propositadas, inteligentes que eles executam, e às operações em que eles se engajam, há algo

mais, uma entidade não-física chamada mente, existindo além do corpo, que é o verdadeiro e oculto repositório do repertório intelectual e mental de uma pessoa.

De acordo com isso, escreve Ryle, em uma passagem que é bem típica: "Os desempenhos observáveis inteligentes não são indícios sobre o mecanismo das mentes; eles são tal mecanismo. Boswell descreveu a mente de Johnson quando descreveu como ele escreveu, falou, comeu, irritou-se e fumou".<sup>15</sup> Em outras palavras, para Ryle, o comportamento de Johnson não é uma evidência de que algo está ocorrendo por trás dele, nos退iros secretos da mente de Johnson. Em vez disso, os processos mentais de Johnson são manifestados em seu comportamento. Seus desempenhos observáveis inteligentes não são um substituto dos mecanismos de sua mente, um mero representante ou procurador. Em vez disso, eles são aqueles mecanismos, completamente abertos para visão e inspeção públicas.

Assim, não é difícil apreciar por que Ryle deveria ter sido rotulado como um behaviorista. Se os processos mentais não são processos fantasmagoricamente escondidos em uma alma imaterial, e se os mecanismos da mente de uma pessoa forem identificados com desempenhos observáveis, públicos, e se a redução da mente à matéria também tiver sido repudiada, é difícil ver no que mais poderiam consistir os fatos sobre a vida mental de uma pessoa, se não em fatos sobre seu comportamento. (Outra opção seria o funcionalismo – ver Capítulo 5 –, mas essa opção não estava disponível a Ryle, uma vez que a teoria ainda não tinha sido inventada no período em que ele estava escrevendo.)

Observe, contudo, que o behaviorismo de Ryle está a uma grande distância do ideal de Hempel de reformular todas as descrições de comportamento em termos dos movimentos corporais incolores

e dos conceitos da física. Ryle claramente não está nem um pouco preocupado com o fato de que sua análise será acusada de circularidade, porque ele não está tentando reduzir os estados mentais a descrições puramente físicas, à maneira dos positivistas lógicos. É por isso que a variedade de behaviorismo de Ryle, em contraste com o behaviorismo "duro" de Hempel, foi rotulada de behaviorismo "brando".<sup>16</sup>

### Ryle e as disposições

Pouco foi dito até o momento sobre a necessidade de se incluírem descrições não apenas do comportamento real, mas também do potencial, em qualquer proposta de análise comportamental dos estados mentais, e é a isso que nos voltaremos agora. Seria altamente implausível sustentar a tese de que os estados mentais consistem inteiramente em comportamento público real e visível, porque isso seria incapaz de acomodar o fato comum de que pessoas possam se encontrar em vários tipos de estados mentais sem jamais os revelar por meio de seus comportamentos. Por exemplo, neste momento eu abrigo vários tipos de crenças e intenções, mas alguém que esteja me observando não poderia dizer exatamente que crenças e intenções são essas. Mesmo se eu tivesse uma dor de dente, isso não seria necessariamente revelado em meu comportamento. Por exemplo, se estou em um concerto, preocupado em não estragar o prazer dos outros, posso muito bem sentar-me e sofrer em silêncio, de modo que ninguém possa descobrir, a partir da minha compostura, as agoniias por que estou passando.

Para evitar essa dificuldade, foi invocada a noção de uma disposição comportamental. Isso quer dizer que a análise comportamental precisa incluir não meramente o comportamento real, mas também o comportamento potencial. Para que al-

guém me descreva verdadeiramente como tendo uma dor de dente, não é necessário que eu esteja realmente gemendo, apertando minha bochecha, e assim por diante. Em vez disso, é suficiente que eu esteja disposto a gemer, a apertar minha bochecha, etc., nas circunstâncias apropriadas. Em outras palavras, isso quer dizer que, se certas condições fossem satisfeitas, então eu gemeria. O mesmo se aplica a estados intencionais, tais como desejos e crenças. Descrever-me como acreditando que a lua é feita de queijo fresco não é, de maneira improvável, comprometer-se a dizer que estou realmente expressando essa crença, mas apenas que, dadas as circunstâncias apropriadas, a expressaria.

A maneira de Ryle estabelecer esse argumento consiste em dizer que ter uma crença é ter uma tendência ou estar propenso a agir de certo modo, e ele procede explicando isso por meio do uso dos exemplos da fragilidade e da solidez. Dizer que um vidro, por exemplo, é frágil não é dizer que ele está realmente se despedaçando, mas apenas que, se ele fosse golpeado, então se despedaçaria. Similarmente, descrever o açúcar como soltível é dizer que, se ele fosse imerso em água, então se dissolveria. De acordo com Ryle, ter uma disposição por parte de algo é simplesmente ser verdadeiro de muitos enunciados hipotéticos "se... então...". Na mesma linha de sua negação de que estados mentais sejam eventos de espécie diferente, e ocultos por trás do comportamento, Ryle está ansioso para enfatizar que ter uma disposição não é encontrá-la em uma certa espécie de estado: "Possuir uma propriedade disposicional não é encontrar-se em um estado particular, ou sofrer uma mudança particular; é estar obrigado ou sujeito a encontrar-se em um estado particular, ou a sofrer uma mudança particular, quando uma condição particular estiver satisfeita."<sup>17</sup> Antes

de prosseguir a leitura, examine o Exercício 4.5.

### **EXERCÍCIO 4.5**

Quão plausível lhe parece a alegação de Ryle de que ter uma disposição consiste simplesmente em um conjunto hipotético de "se ... então ..." ser verdadeiro, desde que as condições apropriadas estejam satisfeitas? Considere sua resposta e discuta-a com outra pessoa antes de ler meus comentários.

A maioria dos filósofos concordaria, hoje em dia, que uma disposição é mais do que Ryle alega que ela seja. A razão pela qual um vidro possui uma disposição para despedaçar-se quando golpeado – a explicação para sua fragilidade – deve ser encontrada nos fatos acerca de sua microestrutura subjacente. Se a microestrutura viesse a ser alterada de uma maneira apropriada, como quando o vidro é temperado, então sua tendência a despedaçar-se quando golpeado seria muito reduzida, ou mesmo removida, como no caso do vidro blindado. Por paridade de argumento, ter uma crença (ou qualquer outro estado mental) não é meramente estar inclinado a comportar-se de certa maneira, mas encontrar-se em um certo estado interno, que pode, dadas as circunstâncias apropriadas, manifestar-se no comportamento externo. Contudo, uma vez que se tenha concedido a propriedade de pensar em crenças e outros estados mentais como causas internas de comportamento externo, o behaviorismo terá sido abandonado.

Essa crítica contra a recusa de Ryle em aceitar que estados mentais disposicionais consistem em algo mais do que tendências a comportar-se de maneiras particulares leva a outra crítica importante contra o behaviorismo. A essência da

objeção é esta: é um lugar-comum que nós explicamos por que as pessoas exibem o comportamento que exibem em termos de seus estados mentais. Por exemplo, uma dor aguda me faz gritar. Atinar que deixei minhas chaves na porta leva-me a dar a volta e ir para casa. O desejo por cigarros, somado à crença de que a loja do outro lado da rua os vende, explica por que atravesso a rua.

Mas se a dor, o atinar-me e o desejo forem todos expressos em termos puramente comportamentais, então como se poderia apelar a eles com o objetivo de explicar meu comportamento, uma vez que isso equivaleria a tentar explicar uma parcela de comportamento em termos de outra parcela? Isso será equivalente à fútil tarefa de tentar explicar o comportamento por referência a si mesmo. Itens que explicam comportamento não podem, por assim dizer, encontrar-se no mesmo nível que o comportamento que eles são chamados a explicar. Preferencialmente, eles devem ser concebidos como as causas subjacentes do comportamento superficial em questão, isto é, eles mesmos devem ser de natureza não-comportamental.

O argumento foi bem estabelecido por Hilary Putnam, ao escrever sobre o que nós queremos dizer quando atribuímos uma enfermidade, tal como esclerose múltipla, a alguém. É plausível dizer que "Normalmente as pessoas que têm esclerose múltipla apresentam alguns ou todos os seguintes sintomas" é uma verdade necessária (analítica). Em outras palavras, é parte do que significamos por "esclerose múltipla" que ela normalmente suscita certos sintomas. Mas, deve-se enfatizar, isso não significa que falar sobre a enfermidade, a esclerose múltipla, possa ser expresso sem perda de significado, em um falar sobre os sintomas da esclerose múltipla. Isso seria análogo ao erro que os behavioristas analíticos cometem quando

tentam expressar o falar sobre “dor” em termos de comportamento de dor real e potencial. O erro que está na base de ambas essas tentativas é o engano de tentar tornar as causas construções lógicas a partir de seus efeitos. Mas dores e enfermidades são causas de comportamentos e de sintomas, respectivamente, e não podem ser reduzidas a seus efeitos, ou construídas a partir deles.<sup>18</sup>

### Teste de Turing

O mesmo erro pode ser comparativamente encontrado no teste de Turing, que pretende capacitar-nos para decidir se uma máquina possui uma mente ou não. A essência do teste é que se duas coisas assemelham-se muito em seu comportamento, e se dissermos que uma delas possui uma mente, então, para evitar inconsistência, devemos admitir que a outra possui uma mente. O teste é configurado da seguinte maneira. Há duas telas: escondido atrás de uma delas encontra-se um computador digital, e atrás da outra, um ser humano. Um pesquisador pode comunicar-se com ambos os sujeitos, mas ele não sabe qual das telas esconde a pessoa e qual esconde a máquina. As respostas fornecidas pelos sujeitos são apresentadas em uma tela. A ideia é que, a partir das respostas fornecidas, o pesquisador deverá dizer qual veio da pessoa e qual, do computador. Se ele não puder notar a diferença, e se admite-se que a pessoa possui uma mente e é consciente, então parece que também se deveria dizer que o computador tem mente e consciência.

Claro que o teste, como foi descrito, é extremamente limitado, porque apenas uma diminuta fração do comportamento, real ou aparente – o comportamento linguístico –, é observada. Assim, empreguemos um pouco de nossa imaginação e

suponhamos que foi criado um androide, algo como o garoto artificial no filme *Inteligência Artificial*, ou Data, na série de televisão *Jornada nas Estrelas*. À parte do fato de que o garoto e Data foram construídos a partir de materiais inorgânicos, seus comportamentos, notadamente incluindo seus comportamentos linguísticos, assemelham-se tanto àqueles de pessoas de carne e osso que qualquer um não ciente de seu *status* de androides pensaria que eles eram humanos. Mesmo que não pudéssemos explicar como tais máquinas fossem conscientes, não deveríamos ser culpados de um preconceito, uma espécie de racismo contra robôs, caso nos negássemos a lhes atribuir estados mentais e consciência?

Mas, embora seja verdade que seu comportamento fornece evidência para a atribuição de estados mentais, não estamos logicamente compelidos a aceitar que androides artificialmente criados possuem mentes. Essa conclusão apenas parecerá inevitável se o comportamento for tratado não meramente como evidência da mentalidade, mas como constitutivo dela, isto é, como constituindo o próprio significado do que seja possuir uma vida mental. O que significa dizer que algo possua uma vida mental reduz-se a como nós constatamos que ele tenha uma vida mental. Isso, como comenta apropriadamente Colin McGinn, é “verificacionismo fora do lugar: ele confunde as questões ontológicas com as epistemológicas”.<sup>19</sup> Também deve ser imediatamente evidente que concluir, com base no teste de Turing, que algo deve possuir uma mente é, com efeito, adotar o behaviorismo viajando por uma rota ligeiramente diferente. O ponto é que algo pode comportar-se como um ser humano e ainda assim ser totalmente desprovido de mente; com efeito, deve ser lembrado, esse é o ponto de partida para o problema das outras mentes.

### A negação dos traços "internos" subjetivos dos estados mentais

Uma das objeções mais fortes contra o behaviorismo deriva-se do senso comum ultrajado. Por causa de sua insistência em reduzir estados mentais a padrões de comportamento externo, ou a disposições para engajar-se em tal comportamento, o behaviorismo efetivamente nega a existência dos aspectos "internos" dos estados mentais, seus qualia. Mas, obviamente, ter dor de dente é mais do que gemer e apertar a própria bochecha, ou mesmo estar disposto a fazer isso. Quando alguém vai a um dentista ou toma uma aspirina, o objetivo não é meramente evitar comportar-se, ou estar disposto a comportar-se, de certa maneira, mas livrar-se da sensação desagradável, da dor que está sentindo. É por essa razão que os behavioristas foram acusados de fingir anestesia. Explicar sensações em termos comportamentais envolve uma dificuldade particularmente intratável, porque é parte da essência das sensações que elas possuam uma fenomenologia. Em resumo, há algo que é como sentir que se tem uma coceira ou um formigamento no pé, ou uma dor no ombro. Tudo isso envolve consciência: há algo que é como passar por tais experiências.

A experiência de visualizar algo com o "olho da mente" também é peculiarmente difícil de explicar em termos comportamentais. Nós pensamos sem hesitação nessa experiência como consistindo em imagens mentais privadas, que frequentemente não se manifestam, e nem precisam manifestar-se, no comportamento externo. Mas, em sua ansiedade para repudiar a existência de estados internos, Ryle escreve:

Uma pessoa imaginando sua estufa é, de uma certa maneira, similar àquela que está vendo sua estufa, mas a similaridade não consiste em seu olhar para uma representação real da estufa, e sim em seu parecer ver sua própria estufa, quando não a está vendo realmente. Ele não está sendo um espectador de uma semelhança de sua estufa, mas se está assemelhando a um espectador de sua estufa.<sup>20</sup>

Antes de ler minha resposta à explanação de Ryle, examine o Exercício 4.6.

### EXERCÍCIOS 4.6

Quão plausível lhe parece a explicação de Ryle sobre o que é para alguém imaginar sua própria estufa? Que críticas você pode apresentar contra essa explicação?

O que Ryle está sustentando é tênue e inconvincente. Para começar, como poderia me assemelhar a um espectador de minha estufa se eu estivesse deitado na cama, nu, no escuro, envolvido em um momento de nostalgia, visualizando minha estufa? Além disso, como já observei, a experiência de visualizar minha estufa é um exemplo primário do tipo de caso em que gostaríamos de dizer que um episódio privado de consciência, não capturável em termos de comportamento público externo, está integralmente envolvido.

### A negação do conhecimento e da autoridade de primeira pessoa relativamente aos estados mentais

Nós concedemos no primeiro capítulo (seção "Avaliando as doutrinas do acesso privilegiado e da translucidez") que há ocasiões nas quais parece que outra

conhecem nossas próprias mentes melhor do que nós mesmos, o que exige que adotemos temporariamente uma perspectiva de terceira pessoa a partir da qual possamos avaliar nosso estado mental. Mas, na maior parte do tempo, e especialmente no que toca a estados não-intencionais, como sensações, nos encontramos na melhor posição para dizer o que estamos sentindo ou pensando. Além disso, podemos descrever nossas sensações, ou expressar nossas crenças ou intenções, sem ter que passar por nenhum processo de investigar quais sejam elas por meio de observação ou inferência a partir de dados mais básicos. Em particular, não preciso observar meu próprio comportamento em um espelho ou vídeo para ser capaz de dizer em que acredito ou como me sinto. O behaviorismo, ao fazer de meus estados mentais uma questão de como me comporto, inverte isso tudo, e leva a piadas, tais como "Um behaviorista encontra outro na rua. Você se sente bem, ele diz ao outro. Como eu me sinto?". Ou um behaviorista diz, depois de fazer amor com outra behaviorista: "Foi muito bom para você. Como foi para mim?". A isso podemos somar a observação ulterior de que, se ter dor ou ter uma crença não fosse nada mais do que comportar-se ou estar disposto a comportar-se de uma certa maneira, obteríamos a consequência absurda de que uma pessoa poderia ter que esperar até ter exibido o comportamento apropriado antes de poder relatar seu estado mental.

### A possibilidade de fingimento pode mostrar que o behaviorismo não pode estar correto?

Pode-se afirmar que o behaviorismo não pode estar correto porque uma pessoa poderia estar meramente fingindo que sen-

te dor. Isso quer dizer que poderia haver abundância de comportamento, mas nenhuma dor correspondente. Como, então, poderia a dor ser idêntica ao comportamento de dor, uma vez que um pode ocorrer sem o outro? O behaviorista, contudo, tem uma resposta. O objetivo do exemplo é mostrar que pode haver comportamento sem um estado mental correspondente – neste caso, dor – e, por conseguinte, estados mentais devem ser mais do que meramente comportamento. Contudo, embora existam boas razões independentes para se pensar que essa conclusão é verdadeira, não parece que ela possa ser estabelecida com base na possibilidade de fingimento. O fingimento implica a presença de uma intenção de fingir, e essa intenção, poderia argumentar o behaviorista, pode ser analisada como o comportamento de dor simulado que constitui o fingimento. Dizendo de outra maneira, é verdade que o comportamento de dor não pode ser identificado com a dor, pois a pessoa não está com dor alguma. Mas, pode o behaviorista sustentar, o comportamento de dor ainda pode ser identificado com *algum* estado mental, a saber, o estado mental mais complexo de pretender fingir. Portanto, o argumento a partir da possibilidade de fingimento não pode mostrar aquilo que pretendia mostrar: que, a despeito da abundância de comportamento, nenhum estado mental está presente, e, assim, o comportamento não pode constituir estados mentais, ao contrário do que sustenta o behaviorismo.

Entretanto, a possibilidade de zumbis, que foi discutida brevemente no Capítulo 3, seção "O argumento de Kripke contra a teoria da identidade", é suficiente para estabelecer a falsidade do behaviorismo. Os zumbis são fisicamente idênticos aos seres humanos, exceto pelo fato de que lhes falta uma vida mental. Para todos os efeitos práticos, eles se comportam exata-

mente como seres humanos, mas lhes falta qualquer consciência ou conhecimento dos movimentos que seus corpos estão efetuando. Em outras palavras, os zumbis exibem todos os tipos adequados de comportamento, tanto quanto um observador externo possa identificar; no entanto, a mentalidade está completamente ausente. Isso parece ser uma fantasia inteligível, não viciada por contradição interna. Sendo isso verdadeiro, então essa razão já seria suficiente para que o behaviorismo fosse rejeitado.<sup>21</sup>

A falsidade do behaviorismo também poderia ser demonstrada pela possibilidade de contrária, a saber, a presença de mentalidade na ausência de quaisquer manifestações comportamentais. Hilary Putnam criou o seguinte experimento mental para apresentar um caso convincente em favor dessa possibilidade. Nós devemos imaginar uma raça de superespartanos ou superrestoicos que constituam uma comunidade na qual os adultos possuem a habilidade de suprimir com sucesso todos os comportamentos de dor involuntários. Embora eles às vezes admitam que estão com dor, eles sempre fazem isso com vozes calmas, bem moduladas, mesmo que estejam sofrendo os tormentos dos condenados. Eles também não gemem, gritam, estremecem, suam ou cerram os dentes. Eles admitem que é necessária uma grande força de vontade para fazer isso, mas têm importantes razões ideológicas para comportar-se do modo como se comportam, e passam por anos de treinamento para alcançar o padrão correto de comportamento frente à dor. Poderia ser argumentado que a ausência de comportamento de dor só será encontrada nos adultos que foram adequadamente condicionados, e que os não-adultos exibirão reações de dor não condicionadas, o que fornece um ponto de apoio para a atribuição de estados mentais à comunidade. No entanto, em resposta a

essa objeção, Putnam incrementa ainda mais a fantasia. Suponhamos que, depois de muitos milhões de anos, os superespartanos começaram a ter crianças que nasciam completamente aculturadas, falando a linguagem dos adultos, compartilhando opiniões sobre sua sociedade e também as crenças sobre não dar sinais de que se está com dor, a menos que seja por meio de um relato verbal, de uma maneira calma, sem ansiedade ou preocupação. Não há, então, respostas não condicionadas de dor nessa comunidade, e, contudo, seria loucura adotar a posição segundo a qual, porque há uma total ausência de comportamento de dor natural, não condicionado, os superespartanos não podem ter a capacidade de experenciar dor.

Putnam também nos pede para imaginarmos super-superespartanos que habitam o mundo-X. Essas pessoas foram superespartanas por tanto tempo que elas não apenas não evidenciam o comportamento de dor não-condicionado, como também começaram a suprimir todo o discurso sobre dor. Habitantes do mundo-X não admitem que têm dores, e fingem desconhecer quer a palavra "dor", quer a experiência à qual ela se refere. Todavia, eles têm dores e sabem que têm. Putnam conclui:

Se esta última fantasia não é, de alguma maneira disfarçada, autocontraditória, então o behaviorismo lógico é simplesmente um erro. Não apenas a segunda tese do behaviorismo lógico – a existência de uma quase-expressão da dor para o discurso sobre comportamento – é falsa, mas até mesmo a primeira tese – a existência de "implicações analíticas". As dores são responsáveis por certos tipos de comportamento – mas apenas no contexto de nossas crenças, desejos, atitudes ideológicas e assim por diante. A partir do simples enunciado "X tem uma dor" não se segue nenhum enunciado comportamental – nem mesmo um enunciado

comportamental contendo um "normalmente" ou um "provavelmente".<sup>22</sup>

O exemplo mais extremo de estados mentais sem comportamento é fornecido pela possibilidade de existência desencarnada. Pode-se argumentar que apenas uma gama limitada de experiências poderia ser atribuída a pessoas desencarnadas, uma vez que elas não teriam corpos nos quais pudessem sentir dores, ou olhos com os quais pudessem ver. Mas não há razão pela qual não devesse parecer para tais pessoas como se elas tivessem corpos e olhos, ainda que não os possuíssem, e que lhes parecesse como se elas estivessem sentindo dores nesses corpos imaginários, e tendo experiências visuais como resultado das operações em seus olhos imaginários. É questionável que o argumento de Descartes a partir das percepções claras e distintas (ver Capítulo 2), assim como o argumento de Kripke contra a teoria da identidade (ver Capítulo 3), seja capaz de estabelecer a possibilidade de estados mentais desencarnados – isto é, que mesmo que tais estados não possam existir neste mundo, há um mundo possível no qual eles existem, e isso é tudo que é necessário para provar que o behaviorismo analítico é falso.

### **Questões para reflexão**

1. Em sua opinião, o behaviorismo apresenta uma solução clara e satisfatória para o problema mente/corpo?
2. O behaviorismo pode explicar a existência de uma vida mental interna na ausência de manifestações comportamentais?

3. Poderiam existir superespartanos que sentem dores, mas nunca exibem comportamento de dor?
4. Os zumbis representam uma possibilidade genuína?
5. Quando digo "Eu tenho uma coceira", estou descrevendo algo que é privado a mim mesmo?

### **Sugestões para leituras aprofundadas**

- PRIEST, Stephen, *Theories of the Mind*. London, Penguin, 1991 apresenta uma introdução excelente e clara ao behaviorismo analítico.
- CHURCHLAND, Paul. *Matter and Consciousness*. Cambridge, MA, MIT Press, 1984 disponibiliza uma abordagem ainda mais estimulante.
- KIM, Jaegwon, *Philosophy of Mind*. Boulder, CO, Westview Press, 1996 apresenta uma das melhores discussões; esse livro é muito útil para estudantes iniciantes que desejam aprofundar-se.
- FOSTER, John, *The Immaterial Self: A Defence of the Cartesian Conception of the Mind*. London, Routledge, 1991 apresenta um tratamento do behaviorismo extremamente rigoroso e completo. Não é recomendado para iniciantes, mas é muito útil em um estágio mais avançado.
- SMITH, Peter e JONES, O. R. *The Philosophy of Mind*. Cambridge, CUP, 1986 fornece uma discussão estimulante e clara do behaviorismo; esse livro é altamente recomendável, especialmente para iniciantes.

## NOTAS

1. Para uma formulação bastante precisa e rigorosa do behaviorismo analítico ver: John Foster, *The Immortal Self: A Defence of the Cartesian Conception of the Mind* (London, Routledge, 1991), p. 33-46.
2. Ver: MALCOLM, Norman, *Dreaming*. London, Routledge and Kegan Paul, 1959. Malcolm contestou que estivesse alegando que sonhos não são nada além de experiências despertas - por exemplo, ele escreve: "De fato eu não estou tentando dizer o que é sonhar; não entendo o que significaria fazer isso", e "não é fácil compreender a relação entre sonhos e convicções despertas de ter sonhado. O sonho e a convicção desperta não são uma e a mesma coisa, no sentido em que a estrela da manhã e a estrela vespertina são uma e a mesma" (p. 59). Não obstante, o veredito filosófico geral parece ter ido contra Malcolm, e muitos filósofos sentiram que a posição que ele estava defendendo equivalia à redução dos sonhos às impressões despertas.
3. A primeira análise da causação de Hume pode ser encontrada em SELBY-BIGGE, L. A. e NIDDITCH, P. H., eds. *Enquiry concerning Human Understanding and concerning the Principles of Morals*, 3<sup>a</sup> ed., Oxford, Oxford University Press, 1995 na seção 7 da primeira investigação sobre o entendimento humano. Posteriormente Hume melhorou essa análise, reconhecendo que, embora o conhecimento de causa e efeito pudesse ser baseado na experiência de regularidades, ele não precisava ser reduzido a meras regularidades. A relação causal é mais bem capturada, Hume reconheceu, pela fórmula "se o primeiro objeto não tivesse existido, o segundo jamais teria existido". Esse enunciado é o que os filósofos chamam de um "condicional contrário aos fatos", e não é equivalente à alegação de que, sempre que um tipo de evento ocorra, ele será seguido por outro tipo. Em vez disso, ele enuncia uma dependência de um evento sobre o outro, dizendo que, se um tipo de evento não ocorresse, então o outro tipo também não ocorreria.
4. Isso lembra SACKS, Oliver, *The Man who Mistook his Wife for a Hat*. London, Picador, 1986, em que relata o caso de um homem que sofria de agnosia visual, e parecia confinado a ser capaz de produzir tais descrições. Ele não podia ver uma luva como uma luva e descrevê-la em termos humanos, como um artigo de vestuário projetado para ajustar-se a membros humanos e satisfazer propósitos humanos, mas podia apenas caracterizá-la científicamente, como "uma superfície contínua, revestida sobre si mesma. Ela parece ter cinco evaginações, se é que essa é a palavra" (p. 13).
5. HEMPEL, C., "The Logical Analysis of Psychology", reimpresso em: BLOCK, Ned, ed., *Readings in the Philosophy of Psychology*, v.1. Cambridge, MA, Harvard University Press, 1980, p. 19.
6. Ibid., p. 17.
7. O proferimento de uma frase não é necessariamente uma afirmação ou um enunciado de algum fato, que pretende transmitir informação, nesse caso, o que se passa com aquela pessoa, do ponto de vista da sua experiência. "Eu tenho uma dor de dente" poderia ser produzido, como de fato está sendo agora, não para informar como eu me sinto, mas meramente para fornecer um exemplo filosófico. Por uma razão diferente, se uma pessoa produzisse as palavras "Eu tenho dor de dente" enquanto está dormindo, isso não poderia ser interpretado como um relato de uma dor de dente, porque a pessoa não tem consciência de que as palavras estão saindo de sua boca, e não tem nenhuma intenção operante de proferir aquelas palavras.
8. FOSTER, John, *The Immortal Self*, p. 29.
9. Ver: MADELL, Geoffrey, *Mind and Materialism*. Edinburgh, Edinburgh University Press, 1988, p. 37-8.
10. Ver: FOSTER, John, *The Immortal Self*, p. 42-3:  
E assim, ao tentar explicitar o significado de "Smith acredita que D fica ao norte", ele [o behaviorista analítico] estará reincidindo em frases da forma:
- 3) Se Smith tivesse, adicionalmente, tais e tais outros estados mentais, ele estaria comportamentalmente disposto a se comportar de tais e tais maneiras, e assim estaria dependendo de uma gama ulterior de conceitos psicológicos cujas análises comportamentais ainda precisam ser fornecidas. Mas então se torna claro que, se estruturado dessa maneira, o processo analítico será interminável, e, de fato, até quando começar a voltar-se sobre si mesmo, cada vez maior. O behaviorista jamais será capaz de alcançar uma interpretação final do enunciado de desejo original, dado que cada vez que ele volta sua atenção para um determinado estado mental que ele até o momento pressupõe, ele será forçado a pressupor outros estados mentais ao especificar sua significância comportamental.
11. SHOEMAKER, S. e SWINBURNE, R. *Personal Identity*. Oxford, Blackwell, 1984, p. 99.

12. RYLE, G., *The Concept of Mind*. Harmondsworth, Penguin, 1973. Ver a observação de Alan Donagan de que *The Concept of Mind* “é um clássico do aristotelismo analítico moderno, vestido de um behaviorismo que agora está fora de moda”, em: *Choice: The Essential Element in Human Action* London, Routledge and Kegan Paul, 1987, p. ix.
13. RYLE, G. *The Concept of Mind*, p. 23.
14. Ibid., p. 17.
15. Ibid., p. 57 (ênfase minha).
16. A relutância de Ryle de engajar-se na espécie de análise que Hempel almejava é ilustrada em sua observação de que *The Concept of Mind* seria estigmatizado, inofensivamente, como behaviorista. Também é relatado que ele teria dito não passar de “behaviorista apenas de um braço e de uma perna” (ver: BIRNE, Alex, “Behaviourism”, in: GUTTENPLAN, S. ed., *A Companion to the Philosophy of Mind*. Oxford, Blackwell, 1994, p. 135).
17. RYLE, G. *The Concept of Mind*, p. 43.
18. PUTNAM, H. “Brains and Behaviour”, in: PUTNAM, H., ed., *Mind, Language and Reality: Philosophical Papers*, v. 2. Cambridge, CUP, 1975, p. 330.
19. MCGINN, Colin, *The Character of Mind*, 2<sup>a</sup> ed. Oxford, OUP, 1996, p. 115.
20. RYLE, G., *The Concept of Mind*, p. 234.
21. Compare isso com a fantasia de Peacocke, de uma marionete, externamente idêntica a um ser humano dotado de uma mente, mas na verdade controlada por rádio a partir de Marte, e completamente desprovida de estados mentais: PEACOCKE, Christopher, *Sense and Content* (Oxford, Oxford University Press, 1983).
22. PUTNAM, H. “Brains and Behaviour”, p. 334.

# 5

## FUNCIONALISMO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender o que é função e o que significa caracterizar a mente como função.
- Entender a diferença entre funcionalismo metafísico e psicofuncionalismo.
- Entender como o funcionalismo pode diferenciar o conceito de uma pessoa do conceito de um ser humano, e como a noção de uma pessoa pode ser compreendida.
- Entender o que é o funcionalismo computacional ou da máquina de Turing.
- Entender alguns dos pontos fortes do funcionalismo.
- Entender as seguintes objeções ou argumentos contra o funcionalismo: liberalismo; argumento da mente chinesa; argumento da sala chinesa; argumentos dos qualia invertidos e dos qualia ausentes.
- Entender o que é o funcionalismo homuncular, e as dificuldades que ele enfrenta.
- Entender o argumento de Jackson contra o fisicalismo e ser capaz de avaliá-lo.

## Introdução

Assim como o behaviorismo analítico, o funcionalismo rejeita a noção da mente como uma entidade, uma substância lógica, seja ela pensada como alma ou, alternativamente, como cérebro. O funcionalismo, como o nome implica, concebe a mente como uma função. Mas o que é uma função? Um modo fácil de entender isso é pensar em um termostato. A função de um termostato é regular a temperatura de uma sala ou edifício. Ele toma certo *input* na forma da temperatura ambiente e então, dependendo de como foi programado, produz um dos resultados seguintes: (1) aciona o sistema de aquecimento, porque a sala está fria demais; (2) desliga o sistema de aquecimento, porque a sala está quente demais. Isso está diagramaticamente representado na Figura 5.1. Como o diagrama ajuda a esclarecer, a função de um termostato é o que ele faz, a tarefa que executa ou o papel que desempenha. Se você estivesse procurando um termostato para comprar, a função seria o foco de seu interesse. De que modo o termostato faz o que faz, como sua organização\* interna lhe permite cumprir seu trabalho e do que ele é feito são usualmente questões de interesse secundário para o consumidor comum, se é que são de algum interesse em absoluto. O que de fato acontece dentro do termostato diz respeito aos técnicos especializados, principalmente se ele deixa de funcionar eficientemente ou para completamente de funcionar. A lição que quero extrair desse exemplo é que uma distinção clara pode ser traçada entre:

- a função de uma coisa – isto é, que trabalho ela desempenha;
- que tipo de mecanismo permite à coisa desempenhar sua função.

Logo que você começa a pensar em funções, miríades de exemplos aparecem. Artefatos de várias espécies desempenham funções: ratoeiras, carburadores, telefones e computadores. A biologia humana e animal também gera muitas instâncias de funções: pense nas funções do coração, dos rins e dos olhos. E, por fim, seres humanos desempenham funções tanto singular quanto coletivamente – por exemplo, os papéis de rainha, de primeiro-ministro, do Parlamento, do Judiciário e assim por diante. Em todos esses casos uma distinção clara pode ser traçada entre o trabalho que o artefato, órgão ou pessoa (ou pessoas) realiza(m) e o material ou mecanismos formais que capacitam sua realização. Mais abstratamente, a noção de uma função pode ser ilustrada matematicamente. Considere a função  $y = x^2$ . Diferentes valores da variável  $x$  – o *input* – são transformados em um *output* – o valor de  $y$  – ao serem elevadas ao quadrado. Assim, um *input* de 2 gera um *output* de 4, enquanto um *input* de 12 gera um *output* de 144, e assim por diante. Compare

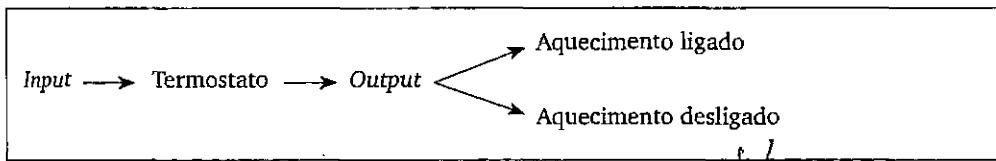


Figura 5.1 Função de um termostato.

\*N. de T: No original, *arrangements*, também traduzido por “arranjos”, “programas de ação”, “mecanismos”, no sentido de “disposição das partes constitutivas”.

também com programas de computador. O conjunto de instruções constituintes de um programa – um algoritmo, como é conhecido – que especifica que passos devem ser dados no processamento da informação também é claramente um exemplo de função. Dados são enviados ao computador que então, de acordo com o algoritmo particular que está sendo usado, transforma-os em um *output*.

A reflexão sobre esses exemplos, à luz da distinção entre a função de uma coisa e os mecanismos que permitem que essa função seja executada, leva à conclusão de que as funções, por suas próprias naturezas, são multiplamente realizáveis. Isto é, um e o mesmo tipo de função pode ser executado por uma variedade indefinida de mecanismos. Podemos dizer que uma e a mesma função pode ser corporificada ou encarnada ou realizada ou instanciada em tipos diferentes de mecanismos. Se o termostato pára de funcionar, eu posso assumir sua função, ainda que seja tedioso sentar em uma sala o dia inteiro ligando e desligando o aquecimento. Se alguém sofre de uma insuficiência cardíaca, um coração mecânico pode ser enxertado para bombear sangue ao resto do corpo. E se a Rainha falece, a monarquia não será imediatamente interrompida, porque o princípio Charles pode rapidamente substituir Sua Majestade.

O ponto ao qual estou querendo chegar é que uma função pode ser especificada abstratamente, e de maneira completamente independente do que quer que habilite a execução dessa função. Claramente, a função não pode ser identificada com os mecanismos materiais que a tornam possível, ou com a pessoa que executa o trabalho, uma vez que esses fatores podem mudar enquanto a função continua sendo executada. Consequentemente, seria errado conceber a função como uma coisa ou processo físico. Mas se a função

não é de natureza física, isso quer dizer que ela é não-física assim como a alma cartesiana? Será ela uma substância-alma fantasmagórica? Eu sugiro que desistamos de pensar em funções dessa maneira, a fim de que não superpovemos o universo com entidades abstratas e também não geremos o pseudoproblema de como as funções se relacionam e interagem com coisas físicas. A melhor concepção de uma função não é nem física, nem não-física. Ela resiste à classificação em ambas as categorias, o que é uma conclusão que, como veremos, harmoniza-se com o que muitos filósofos funcionalistas tiveram a dizer sobre a natureza das funções.

Entretanto, embora funções não sejam idênticas aos mecanismos que as corporificam, elas requerem alguma espécie de corporificação para que sejam executadas. A fim de que a função de bombeiar gasolina ao motor ocorra, algum conjunto de programas de ação deve existir para que isso ocorra. Sem uma corporificação de alguma espécie, nada pode acontecer ou acontecerá. A carburação permaneceria meramente uma ideia na mente de seu inventor, algo no papel que na teoria poderia funcionar, mas que não acontecerá até um carburador ser efetivamente construído e instalado.

### Funcionalismo metafísico

O que acontece se tentamos aplicar uma análise funcional a estados mentais, digamos, a um tipo particular de estado mental, uma dor? Suponha que tomemos nossa pista do exemplo do termostato. A função de um termostato é tomar certo *input* e transformá-lo em certo *output*. Bem, talvez a dor também possa ser concebida do mesmo modo. O que tipicamente acontece quando alguém experimenta uma dor é que uma lesão ocorre em algu-

ma parte de seu corpo, a dor é experimentada, e, como resultado, a pessoa se contorce e gême ou esfrega a parte afetada, ao mesmo tempo em que tenta escapar do que quer que esteja causando a dor.

Dito de maneira mais formal em termos funcionais, uma dor pode ser funcionalmente especificada como consistindo em um *input* na forma de lesão de um tecido, um *output* na forma de comportamento de dor – contorção, gemido e assim por diante – e uma relação com outros estados mentais – tipicamente, ela provoca um desejo de se ver livre da dor. Naturalmente, sendo ele próprio um estado mental, o desejo de se ver livre da dor também precisará de uma especificação funcional (ver Figura 5.2). Nesse nível de análise, a especificação do estado mental é altamente abstrata. Uma vez que tenhamos caracterizado o estado mental em termos de *inputs*, *outputs* e relações com outros estados mentais, teremos fornecido uma descrição exaustiva dele: o estado mental, concebido como uma função, não engloba nada mais do que o sistema de relações descritas em termos de *inputs*, *outputs* e outros estados mentais, que por sua vez são também funcionalmente analisados. Esse tipo de funcionalismo, em que o estado mental é especificado de maneira puramente formal, é geralmente conhecido como funcionalismo metafísico, embora seja às vezes chamado de funcionalismo simples. Essa especificação formal é obtida simplesmente por meio de

análise filosófica e tira proveito de nossa compreensão de senso comum do que os estados mentais são. Ela é, supostamente, a expressão de uma verdade conceitual, algo que é necessariamente o caso e que não poderia ser de outro modo.

### Psicofuncionalismo e o conceito de uma pessoa

Contudo, isso deixa em aberto a questão sobre que conjunto de mecanismos habilita a função, especificada formalmente, a ser executada, e isso não pode ser obtido a partir do conhecimento da função tomada isoladamente. É perfeitamente possível perceber o que termostatos, corações e monarcas fazem sem se ter a menor ideia de como termostatos e corações são construídos, ou quem é o monarca efetivamente. Do ponto de vista do funcionalista metafísico, a mente é uma caixa preta, como os psicólogos a chamam, mediando *inputs* e *outputs*, mas opaca no que concerne ao que efetivamente ocorre em seu interior. Ela funciona bem, mas como ela funciona é algo desconhecido.

Embora existam funcionalistas que, em teoria, estejam dispostos a conceder que o que está na caixa preta é algo não-físico no formato de uma alma fantasmagórica, cuja função é executada por processos dentro de suas entranhas etéreas, na prática todos os funcionalistas são materialistas ou fisicalistas, que sustentam

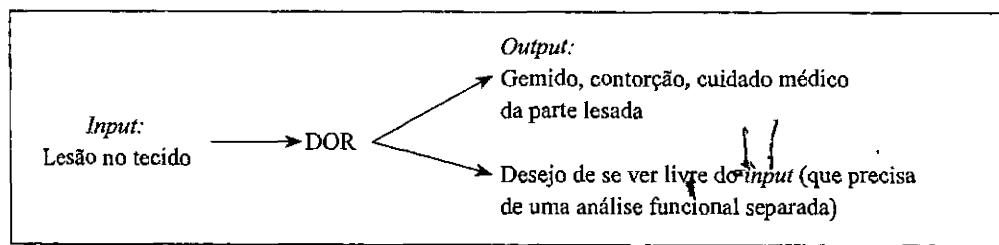


Figura 5.2 Especificação funcional da dor.

que as funções mentais são encarnadas em processos neurofisiológicos no cérebro e sistema nervoso central. Eles argumentam que, uma vez que a noção de uma alma é altamente dúbia (ver Capítulo 2, seção “O problema de contar almas”), apenas a segunda alternativa materialista é uma opção realista. Contudo, uma vez que a ideia de almas não-físicas faça sentido, a primeira opção representa uma possibilidade genuína, porque o funcionalismo metafísico é neutro em relação a uma metafísica materialista ou imaterialista. A função ela mesma, como eu estava com dificuldades para enfatizar antes, não pode ser classificada nem como não-física nem como física, e nunca pode ser estritamente identificada com o sistema de mecanismos que efetivamente a instantiam, seja qual for a natureza final que se prove que esses mecanismos possuem. De fato, a neutralidade do funcionalismo metafísico foi percebida como uma vantagem pelos materialistas, precisamente porque ela não impede a possibilidade de que uma função possa ser executada por um conjunto puramente material de mecanismos.

Quantidade alguma de teorização filosófica ou especulação sozinha revelará que mecanismos neurofisiológicos são responsáveis por corporificar dor como um estado funcional. Isso é um assunto que requer investigação empírica. Este é

o projeto do psicofuncionalismo, a tarefa de explicitar os detalhes do que veio a ser chamado de ocupante causal do papel funcional. Nos seres humanos, como dissemos, isso consistirá em processos no sistema nervoso humano. Mas psicofuncionalistas, sem esquecer suas raízes metafísicas, estavam dispostos a ir além e aceitar que é concebível que funções mentais possam ser corporificadas em outras espécies muito diferentes de mecanismo físico.

Assim, o funcionalismo implicitamente reconhece que, embora por questão de fato contingente as únicas pessoas que eu conheço são seres humanos, o conceito de uma pessoa e o conceito de um ser humano não são idênticos no que toca ao significado. Não é logicamente necessário que uma pessoa seja um ser humano, porque em teoria poderiam existir pessoas não-humanas – marcianos, por exemplo – cujas cabeças contêm a gosma verde tradicional, que corporifica as funções mentais que nos seres humanos são executadas pelo cérebro e pelo sistema nervoso central. Conceivelmente, se há androides como Data em *Jornada nas Estrelas*, então a mentalidade pode ser corporificada nos *chips* que constituem um cérebro de silício. O diagrama na Figura 5.3 resume isso. Ser um ser humano tampouco é suficiente para o *status* de pessoa, porque há seres humanos que deixaram

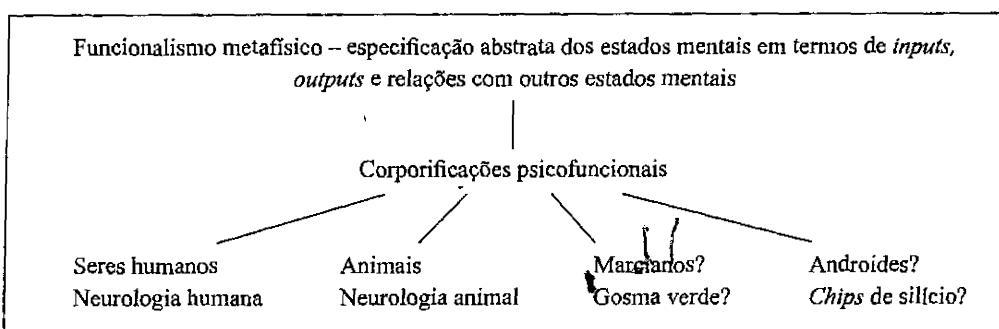


Figura 5.3 Funcionalismo metafísico e psicofuncionalismo.

de ser gente, ou pessoas, no sentido mais completo, tais como aqueles que tiveram seus cérebros irreversível e severamente lesados, tendo sido reduzidos a um estado semivegetativo. E há seres humanos que jamais conseguirão se desenvolver propriamente como pessoas, porque alguma doença ou trauma lesou seus sistemas nervosos enquanto eles estavam se desenvolvendo no útero.

Essas observações levantam a questão sobre como o conceito de uma pessoa, por oposição ao conceito de um ser humano, deve ser entendido. A classificação de alguém como *Homo sapiens* é uma classificação biológica, e os detalhes exatos do que deve ser verdadeiro acerca de um organismo para que ele seja considerado humano é uma tarefa para os biólogos explicarem em detalhes. Mas tentar dizer o que é característico de pessoas exige análise filosófica. Assim como ocorre com muitos outros conceitos, é muito improvável que uma exposição de condições necessárias e suficientes para que alguém se qualifique como pessoa possa ser fornecida, e as fronteiras dos conceitos raramente são bem definidas, de modo que as observações a seguir não têm a intenção de ser rigidamente definitivas.

O que, então, é distintivo de pessoas por oposição a animais não-humanos que não são admitidos na classe de pessoas?<sup>1</sup> A resposta não pode ser a posse de consciência porque, contra Descartes, não há razão para negar que animais, pássaros, peixes e talvez até mesmo insetos estejam cientes do mundo que habitam. Descartes criou um abismo entre nós e os animais não-humanos, sustentando que eles eram meros autômatos mecânicos, desprovidos de qualquer vestígio de sensação, consciência ou pensamento; porém, um quadro mais acurado é fornecido por uma concepção aristotélica da mente como uma hierarquia emergente de capacida-

des que aumentam em complexidade e sofisticação à medida que ascendemos na árvore evolutiva (ver Capítulo 8, seção “*Psiche* e a hierarquia das capacidades”).

Os animais são capazes de experimentar uma ampla gama de estados mentais: sensações, percepções, emoções primitivas tais como aversão e medo, a compulsão dos apetites e necessidades, e mesmo, no caso dos animais superiores, crenças. Um cão latindo em frente a uma árvore desabitada, que costumava abrigar um esquilo, deve, aparentemente, em algum sentido acreditar que o objeto de suas atenções continua na árvore, embora não possa ser creditado com uma crença na verdade da proposição de que há um esquilo na árvore, uma vez que crenças sofisticadas dessa espécie exigem a capacidade para aquisição e uso da linguagem. Do nosso ponto de vista como usuários da linguagem, não podemos especificar o conteúdo da crença do cão senão por meio de uma forma proposicional usando uma cláusula “que”, mas o conteúdo intencional da mente do cão não consiste em entreter e comprometer-se com a aceitação da verdade de uma proposição.<sup>2</sup> Isso significa que é difícil explicar em que consiste o conteúdo intencional no caso de animais: como P. F. Strawson observou em outro contexto, nos faltam as palavras para poder dizer como é estar sem elas.<sup>3</sup>

Uma resposta mais promissora para nossa questão é que as pessoas são dotadas da capacidade para autoconsciência. Elas podem formar um conceito de suas próprias identidades ao longo do tempo. Podem pensar em si mesmas e se autorreferir como uma e a mesma pessoa que elas são agora, foram em algum tempo no passado, e serão no futuro. Elas possuem uma concepção de suas origens, das circunstâncias e dos locais de seus nascimentos, de suas famílias e da sociedade mais ampla que as circunda, de seu provável



progresso pelo mundo, e do inevitável e avultante fato de suas próprias mortes. As pessoas também podem, é claro, pensar no passado e no futuro em geral, mesmo quando eles não mantêm relação especial alguma com suas próprias situações, e também engajam-se em especulação sobre estados de coisas hipotéticos.

Uma pessoa pode atribuir reflexivamente estados mentais a si mesma, assim como especificar sua localização no tempo e no espaço. Ser capaz de fazer isso tudo, de fazer uma referência que identifica a si mesma, e de predicar a si mesma um estado mental particular, exige a posse de uma linguagem contendo o aparato necessário das expressões referenciais e predicativas. Portanto animais, desprovidos de linguagem, estão para sempre impedidos de participar das formas superiores de autoconsciência que caracterizam as pessoas.

Com a linguagem também surge a possibilidade de formular explicações, fornecer razões para crenças e justificações para ações. Uma pessoa pode descrever o que está fazendo ou tentando fazer, explicitar os fins que tem em vista e os meios adotados para assegurar tais fins. Desse modo, pessoas são capazes de avaliar seus desejos, estimando-os de acordo com sua praticabilidade e premênciia, a abandonando-os se necessário à luz de suas estimativas, de modo que eles jamais venham a obter o *status* de intenções firmes e completas. Às vezes expressa-se isso dizendo que pessoas podem formar desejos de segunda ordem – isto é, eles podem querer querer algo ou querer não querer algo. Por exemplo, alguém pode ansiar fumar, mas, reconhecendo os perigos que isso impõe à sua saúde, querer abandonar a ânsia por fumar. Animais, por contraste, não podem fazer nenhuma dessas coisas. Em uma expressão memorável derivada de Harry Frankfurt, animais são *voluptuosos*:

eles agem irrefletidamente e imediatamente quando são estimulados por necessidades e apetites, a não ser que sejam impedidos.

Podemos resumir isso dizendo que pessoas, diferentemente de animais, são capazes de escolhas deliberadas genuínas. Elas podem escolher racionalmente um meio para um determinado fim e avaliar tanto os meios quanto os fins não apenas em termos de viabilidade, mas também moralmente, levando em consideração os prováveis efeitos para o bem-estar e o interesse dos outros, e, se necessário, prescindir de ir adiante. É por isso que um aspecto importante do conceito de uma pessoa é que concebemos as pessoas como centros de responsabilidade moral que podem ser considerados responsáveis e aptos a responder pelo que fazem. É nessa arena que a autonomia e o livre-arbítrio das pessoas devem ser encontrados. À medida que as pessoas se desenvolvem, elas criam suas próprias identidades, pintando o retrato de suas próprias vidas, como diz Sartre. Mas animais, privados de qualquer concepção sobre o que eles são, ou deveriam ser, ou querem ser, não podem ser considerados autônomos nesse sentido. Uma pessoa pode entender o que ela fez e por que isso pode ser considerado por outros como indesejável ou proibido por um código ou uma lei moral cuja violação implicará algum tipo de penalidade ou punição. É por essas razões que o ato de infringir dor ou sofrimento a um animal em resposta a algo que ele fez não pode estritamente ser concebido como punição no sentido em que o usamos com pessoas – punição pressupõe responsabilidade moral, e isso é algo de que os animais estão desprovidos e que sequer podem conceber.

A linguagem estende amplamente a gama de emoções possíveis, porque muitas emoções incorporam pensamentos e

crenças sofisticadas que só são possíveis por meio dela. Um cão não pode achar uma piada sarcástica engraçada, porque o cão é incapaz de compreender o propósito da piada ou seu conteúdo intelectual. De maneira similar, animais não podem sentir remorso, compaixão ou indignação moral, porque sem linguagem eles estão desprovidos de conceitos, valores e crenças morais que tornam possíveis essas emoções complexas.

Relacionamentos entre pessoas, por serem mediados pela linguagem, também são dramaticamente diferentes de relacionamentos entre animais. As pessoas possuem a capacidade de se envolver em debates baseados em raciocínios sobre propriedades ou territórios em disputa, engajando-se em negociação, fazendo concessões e assumindo compromissos. Elas podem fazer promessas e firmar, renegociar ou rejeitar contratos. Uma pessoa tem a capacidade de refletir sobre o bem-estar de outra pessoa, e colocar a si mesma imaginativamente na pele dessa outra pessoa para tentar ver o mundo a partir do ponto de vista dela. Ao reconhecer a injustiça cometida em relação a outra pessoa, alguém pode buscar perdão, e a pessoa injustiçada pode encontrar em si mesma o perdão a ser dado. É dessas maneiras que a gama de simpatias naturais e de compaixão de uma pessoa podem ser estendidas. É quase desnecessário dizer que nenhuma das atitudes ou atividades aqui descritas são encontradas reino animal.

Animais também são incapazes de apreciação estética; eles não podem participar na contemplação arrebatada da serenidade e elegância da seção de abertura das *Variações Goldberg*, de Bach, da

sutileza e solenidade dos "Church Monuments"<sup>\*</sup>, de George Herbert, da imagem angustiante e atormentadora d'*O Grito*, de Munch, ou do sentimento de melancolia que o *Auf der Donau*<sup>\*\*</sup>, de Schubert, inspira. A habilidade do joão-de-barro de criar uma moradia para uma possível parceira é o resultado de um instinto inato que nunca varia e que não é genuinamente criativo. Dando forma à arte criada pelas pessoas, há tradições e movimentos, maneiras de ver o mundo, costumes, convenções e reações a elas, que mudam e evoluem constantemente. Mas disso não encontramos traço algum nos animais.

A posse do *status* de pessoa é tipicamente uma questão que admite graus, não é de tudo ou nada. Uma criança recém-nascida dificilmente seria considerada uma pessoa de acordo com os critérios apresentados anteriormente, mas com o passar do tempo, e dadas as condições adequadas, o bebê gradualmente amadurecerá até tornar-se uma pessoa. É por isso que eles, ao contrário dos animais não-humanos, são admitidos plenamente no prestimoso círculo. No caso dos que sofrem de severas deficiências mentais ou dos insanos, não consideramos tais seres humanos como meras coisas que podem ser tratadas como bem entendemos, mas como também fazendo parte da comunidade de pessoas e como merecedores de nossa consideração, porque, se as coisas tivessem sido diferentes, eles estariam aptos a tomar plena posse de seus lugares ao nosso lado. O que governa nossa atitude é o reconhecimento expresso no dito segundo o qual cabe somente à graça divina decidir quem vai e quem fica.<sup>\*\*\*</sup>

\* N. de T.: Trata-se de um poema.

\*\* N. de T.: Sobre o Danúbio.

\*\*\* N. de T.: O dito original – "There, but for the grace of God, go you and I" – é atribuído a John Bradford, que teria feito essa observação ao ver criminosos sendo levados ao local de execução (c. 1553), daí derivando o sentido de que "eu poderia estar naquela mesma situação terrível".

Voltando brevemente ao funcionalismo, podemos notar que, tão logo as formas de vida demonstrem possuir um grau suficiente dos tipos de capacidades e atitudes aqui esboçadas, independentemente de serem membros da espécie humana, poderíamos inteligivelmente concebê-las como pessoas, e sem dúvida poderíamos achar isso inevitável.

O funcionalismo também está bem posicionado para contornar a objeção que teve um papel central no abandono da teoria da identidade tipo-tipo mente/cérebro, a saber, que é excessivamente implausível insistir que um dado tipo de estado mental seja sempre e em todo lugar idêntico a um dado tipo de estado cerebral. De fato, a teoria funcionalista, parcialmente inspirada pelo surgimento da ciência da computação, foi expressamente projetada para acomodar a objeção e, ao fazê-lo, torná-la nula. Ela pareceria assegurar todas as vantagens do materialismo sem jamais ter que ser estritamente identificada com essa posição, uma vez que funções não são redutíveis a mecanismos materiais que as tornam possíveis, ao contrário do que a teoria da identidade mente/cérebro afirma.

### Funcionalismo computacional ou da máquina de Turing

Os computadores têm um único propósito, o de computar funções. De fato é assim que um computador pode ser definido. É um dispositivo criado para tomar certo *input* e transformá-lo, de acordo com um conjunto de instruções, passo a passo, em um *output*. (Obviamente estou pensando em um computador digital aqui. Deixo de lado os computadores analógicos, tais como redes neurais, que

trabalham em uma base completamente diferente.) Como John Searle tem muitas vezes nos lembrado, a cada era a tecnologia mais recente tem sido tomada como modelo para a mente. Os antigos gregos, relata Searle, pensavam que a mente assemelhava-se a uma catapulta. Leibniz, como vimos (Capítulo 1), concebia a mente como um moinho. No século XX, para o fisiólogo Sir Charles Sherrington (1857-1952), a mente era como uma conversa telefônica. Não é de se admirar, portanto, nessa era da computação e do processamento de informação, que as mentes tenham sido concebidas como dispositivos computacionais naturais que tomam informação na forma de estímulos físicos ou sensoriais e a processem, transformando-a em um *output* comportamental.

Como os computadores computam funções? Um algoritmo é um procedimento, um método para computar uma função. Ele consiste em um número finito de passos distintos, que devem ser tomados para que os valores que compõem o *input* de uma função sejam transformados, passo a passo, no *output* da função. Cada passo a ser tomado é submetido a uma especificação de modo definido, claro e não ambíguo, que não requer *insight* algum ou proeza de imaginação.

Para entender isso mais claramente, considere o seguinte exemplo. Suponha que você está na Inglaterra\* e tira todas as moedas de seus bolsos e posiciona-as aleatoriamente em fileiras, do jeito que elas vão sendo apanhadas. Suponha, além disso, que suas fileiras de moedas sejam como aquelas da Tabela 5.1. Se eu agora lhe pedisse para ordenar as moedas, do maior ao menor valor, você rapidamente pegaria a moeda de £ 2,00, depois a de £ 1,00 que colocaria debaixo da primeira, depois a de £ 0,50 e assim por diante.

\*N. de T.: O exemplo teve de ser mantido como tal, porque o sistema monetário da Inglaterra difere do nosso, com duas moedas a mais.

te. Se eu perguntasse como você fez isso, você provavelmente diria apenas que podia reconhecer o valor de cada uma das moedas e, sabendo disso, ordená-las era uma questão simples. Mas o pedido que fiz para que você ordenasse as moedas da maior à menor não é uma ordem que possamos dar a um computador. Não podemos simplesmente dizer ao computador, como podemos dizer para um ser humano, "coloque as moedas em ordem", e então esperar seriamente que o computador responda. A instrução básica – que obviamente deveria ser dada em uma linguagem de programação – para permitir ao computador realizar a tarefa seria algo como o seguinte. Compare as magnitudes das moedas nas primeiras duas fileiras da tabela. Se a magnitude da moeda na primeira fileira for maior que a da moeda na segunda fileira, deixe ambas as moedas como estão; senão, movimente a moeda na segunda fileira para a primeira fileira. No caso ora em pauta (ver Tabela 5.1), é óbvio que as moedas permanecerão intocadas. Agora repita esse procedimento com as moedas na segunda e na terceira fileiras. Nesse caso a moeda de £ 0,50 será movida para a segunda fileira e a de £ 0,02 será movida para a terceira fileira. Agora compare as moedas na terceira e quarta fileiras e aplique o mesmo proce-

dimento. Faça isso até que as fileiras sete e oito tenham sido comparadas. A Tabela 5.2 mostra a ordem das moedas depois da primeira aplicação desse procedimento.

Claramente, como podemos ver na Tabela 5.2, as moedas ainda não foram ordenadas do maior ao menor valor, mas o computador poderia ser instruído a aplicar o procedimento novamente; e de fato tantas vezes quanto for necessário até que nenhuma alteração ulterior na posição das moedas precise ser feita, quando o programa poderá ser encerrado (de fato, isso exigirá mais três rodadas completas – faça o teste!). Seria incrivelmente tedioso e demorado para um ser humano ter que aplicar esse procedimento toda vez que moedas ou colunas de números precisassem ser ordenadas, digamos, quando passamos as notas dos alunos para uma lista. Mas isso que eu descrevi, com efeito, ainda que de maneira meio simplificada e crua, é um algoritmo, uma série de passos separados, que um dispositivo não dotado de inteligência, de compreensão, ou de uma mente, tal como um computador, poderia usar para transformar valores de moedas em ordem aleatória em uma lista ordenada, que vai do maior ao menor valor de face.

Devo agora introduzir a ideia de uma máquina de Turing, batizada em homen-

Tabela 5.1

| Fileira        | Moeda  |
|----------------|--------|
| 1 <sup>a</sup> | £ 2,00 |
| 2 <sup>a</sup> | £ 0,02 |
| 3 <sup>a</sup> | £ 0,50 |
| 4 <sup>a</sup> | £ 0,20 |
| 5 <sup>a</sup> | £ 0,01 |
| 6 <sup>a</sup> | £ 0,10 |
| 7 <sup>a</sup> | £ 1,00 |
| 8 <sup>a</sup> | £ 0,05 |

Tabela 5.2

| Fileira        | Moeda  |
|----------------|--------|
| 1 <sup>a</sup> | £ 2,00 |
| 2 <sup>a</sup> | £ 0,50 |
| 3 <sup>a</sup> | £ 0,20 |
| 4 <sup>a</sup> | £ 0,02 |
| 5 <sup>a</sup> | £ 0,10 |
| 6 <sup>a</sup> | £ 1,00 |
| 7 <sup>a</sup> | £ 0,05 |
| 8 <sup>a</sup> | £ 0,01 |

gem ao matemático Alan Turing (1912-1954), um gênio da matemática que, entre outros feitos, decifrou o código ENIGMA dos nazistas durante a segunda guerra mundial. Uma máquina de Turing não é uma máquina de verdade, mas um plano para uma máquina possível. Basicamente, tal máquina é uma maneira de explicar como algoritmos podem ser computados. A máquina consiste em uma fita dividida em células e um dispositivo, um cabeçote, que pode escrever um símbolo em uma célula vazia ou reescrever um símbolo preexistente, mas apenas um símbolo em cada célula. O cabeçote pode também ler os símbolos constantes na fita. A fita pode ser movida da direita para a esquerda, e da esquerda para a direita, mas apenas uma célula de cada vez. A máquina também é capaz de encontrar-se em um de dois estados internos, que eu vou chamar E1 e E2. Há apenas quatro coisas que a máquina pode fazer: mover a fita para a direita ou para a esquerda; escrever um símbolo na fita; apagar um símbolo da fita; mudar seu estado interno de E1 para E2, ou de E2 para E1. O que a máquina efetivamente faz quando está em operação é especificado pela sua, assim chamada, tabela de ação. A tabela de ação consiste em uma lista de instruções que especificam o que a máquina deve fazer. A

forma geral das instruções será que, se a máquina encontra-se no estado E1 e está lendo certo símbolo, X, então ela deverá ou escrever ou apagar um símbolo, continuar no estado E1 ou mudar para o estado E2, e mover a fita para a esquerda ou para a direita. Um exemplo simples e perspicuo desse procedimento, que também ilustra como a máquina de Turing pode adicionar 1 a um número existente, é fornecido por Tim Crane em seu excelente livro, *The Mechanical Mind*\*.

Crane nos pede para imaginar que os números possam ser representados por simples barras verticais, tais como as marcas de giz que um prisioneiro poderia fazer em uma parede da cela enquanto conta os dias que passam. Assim ||||| significa 5, || significa 2, | significa 1, e assim por diante. Essas marcas podem ser escritas em uma fita – uma marca por célula –, e diferentes grupos de números podem ser separados por zeros, assim:

...000||||000|000||||||  
representa a sequência de números 5, 1, 7...

Os pontos mostram que a fita estende-se em ambas as direções indefinidamente.

O que precisamos agora é de uma tabela de ação que especifique o que a má-

|                | <i>Input</i>  |   |
|----------------|---|---|
| Estado interno |   | 0   |
| E1             | <ul style="list-style-type: none"> <li>• mudar para estado E2</li> <li>• escrever um  </li> <li>• mover a fita para a direita</li> </ul>    | <ul style="list-style-type: none"> <li>• permanecer no estado E1</li> <li>• escrever um 0</li> <li>• mover a fita para a direita</li> </ul> |
| E2             | <ul style="list-style-type: none"> <li>• permanecer no estado E2</li> <li>• escrever um  </li> <li>• mover a fita para a direita</li> </ul> | <ul style="list-style-type: none"> <li>• mudar para estado E1</li> <li>• escrever um  </li> <li>• parar</li> </ul>                          |

Figura 5.4 Tabela de ação.

\*N. de T.: Em português: *A mente mecânica*.

quina deve fazer para adicionar 1 a um número existente – ver Figura 5.4. Suponha que a máquina seja agora alimentada com uma fita que contém um 2 escrito nela, de modo que ela terá a seguinte apresentação (lembrando que 2 será representado por 2 barras verticais):

0 0 0 | | 0 0 0

A máquina deve desempenhar uma operação muito simples: adicionar 1 a esse número seguindo as instruções dadas na tabela de ação. Ela começa em E1 lendo o primeiro símbolo à extrema direita da fita. A instrução na tabela diz que a máquina deve escrever um 0, continuar em E1, e mover a fita uma célula à direita. Ela faz isso, e, enquanto ainda encontra-se no estado E1, ela encontra outro 0 na coluna próxima à da extrema direita, e assim uma vez mais ela escreve um 0 e move a fita um espaço para a direita. À

medida que a fita se move da esquerda para a direita e a máquina segue as instruções da tabela de ação, o resultado obtido em cada passo é mostrado na Tabela 5.3 (a linha sob os números indica que célula o cabeçote está lendo e ^ representa o cabeçote, que permanece fixo enquanto a fita move-se célula por célula à direita). Em (f), a máquina encontra-se no estado E2 quando encontra uma |. A tabela de ação a instrui a escrever uma |, voltar ao estado E1 e então parar. O que está agora na fita são três e não duas barras verticais, representando o número 3. Desse modo, ao mover a fita uma célula por vez e escrever ou apagar um símbolo, a máquina adicionou 1 a 2.

Adicionar 1 a 2 é uma operação muito simples, mas se essa computação pode ser feita por uma máquina de Turing, então em teoria qualquer função capaz de ser

Tabela 5.3

| Fita            | Operação da máquina   |
|-----------------|---|
| (a) 00011000    | No estado E1, leia 0, grave 0, mova a fita para a direita           |
| (b) 0001100 ^ 0 | No estado E1, leia 0, grave 0, mova a fita para a direita           |
| (c) 000110 ^ 00 | No estado E1, leia 0, grave 0, mova a fita para a direita           |
| (d) 00011 ^ 000 | Mude para estado E2, leia 1, grave 1, mova a fita para a direita    |
| (e) 0001 ^ 1000 | Permaneça no estado E2, leia 1, grave 1, mova a fita para a direita |
| (f) 000 ^ 11000 | No estado E2, leia 0, grave 1, mude para estado E1, pare            |
| (g) 001 ^ 11000 | No estado E1: a leitura na fita mudou, a máquina pára               |

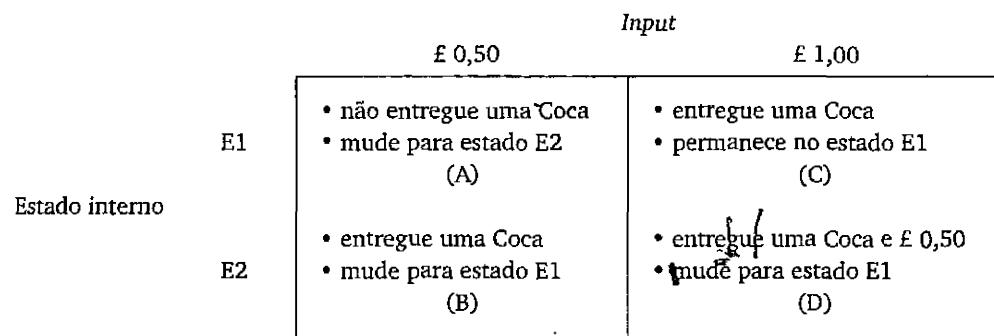


Figura 5.5 Tabela de ação de uma máquina de Coca-Cola.

computada de acordo com um algoritmo (uma série de passos simples, separados) pode ser feita em uma máquina dessas. Obviamente, para executar outras operações, tais como multiplicação, divisão, fatoriais e assim por diante, o conjunto de instruções que compõem a tabela de ação deverá ser mais complexo, mas em princípio não há o que impeça que projetemos tabelas mais complexas. É claro, os números em computadores reais não são representados por séries de barras verticais como no exemplo dado, mas são escritos em código binário: 1 = 1, 2 = 10, 3 = 11, 4 = 100, etc., mas os princípios de operação da máquina de Turing continuam os mesmos.

Tabelas de ação para máquinas de Turing podem ser usadas para desempenhar todo tipo de função, além das puramente matemáticas. Um exemplo (fornecido pelo filósofo Ned Block<sup>5</sup>) que você encontrará com frequência na literatura sobre o funcionalismo é aquele da máquina de Coca-Cola. O que queremos é uma máquina que aceite moedas de £ 0,50 e £ 1,00 e que faça quatro coisas possíveis: (1) faça a entrega de uma garrafa de Coca, se uma moeda de £ 1,00 ou duas moedas de £ 0,50 forem inseridas; (2) não faça nada se apenas uma moeda de £ 0,50 for inserida; (3) faça a entrega de uma Coca se outra moeda de £ 0,50 for adicionada àquela já inserida; faça a entrega de uma Coca e £ 0,50 de troco se £ 1,50 forem inseridos na máquina por engano. A Figura 5.5 ilustra como a tabela de ação se parecerá. A tabela de ação descreve as funções que a máquina desempenha, dependendo de:

1. Seu estado interno – estado E1 ou E2.
2. *Inputs* inseridos na máquina –
  - (a) uma moeda de £ 0,50;
  - (b) uma moeda de £ 1,00 (ou um total de £ 1,00 constituído de duas moedas de £ 0,50);

- (c) uma moeda de £ 0,50 e uma moeda de £ 1,00 (ou £ 1,00 constituído de duas novas moedas de £ 0,50).

Há quatro maneiras possíveis, (A), (B), (C) e (D), nas quais a máquina pode comportar-se dependendo do que um consumidor faça. São elas:

- A) A máquina encontra-se no estado E1. Um consumidor insere £ 0,50. Nada sai da máquina, mas internamente seu estado muda para E2.
- B) A máquina encontra-se no estado E2, sendo que uma moeda de £ 0,50 foi anteriormente inserida. Suponha que o consumidor então insira mais £ 0,50. A máquina faz a entrega de uma Coca e retorna ao estado E1, para aguardar a chegada de um novo consumidor.
- C) A máquina encontra-se no estado E1. Um consumidor insere uma moeda de £ 1,00. A máquina faz a entrega de uma Coca e permanece no estado E1 para aguardar a chegada de um novo consumidor.
- D) A máquina encontra-se no estado E2 porque £ 0,50 acabaram de ser inseridos. Uma moeda de £ 1,00 é então inserida por engano. A máquina faz a entrega de uma Coca e uma moeda de £ 0,50 e então retorna ao estado E1.

O estado funcional da máquina é definido, e de maneira exaustiva, pela descrição dos *inputs*, estados internos e *outputs*, como indicado pela tabela de ação fornecida na Figura 5.5. O estado funcional como é descrito pela tabela de ação é abstrato e puramente formal, o que quer dizer que ele exemplifica o nível de análise do funcionalismo metafísico. Para a função ser executada, ela deve ser corporificada em algum meio que possa instanciar as funções. A natureza do meio realiza-

dor pode, como vimos anteriormente, ser indefinidamente variada em seu caráter. No caso da máquina de Coca ela pode ser constituída de instruções que são corporificadas em um sistema eletrônico, ou na operação de vários dispositivos mecânicos tais como alavancas que abrem e fecham aberturas – de modo que a Coca possa ser liberada por uma canaleta, por exemplo. Ela poderia até mesmo, como diz Block, ser um “objeto disperso” cujas partes estão espalhadas sobre o mundo inteiro e se comunicam via ondas de rádio. Alternativamente, ela poderia ser executada por uma pessoa, ou por um grupo de pessoas, que ficam dentro da máquina e seguem o conjunto de instruções escritas que definem a sua tabela de ação.

Como é que tudo isso se relaciona com o funcionalismo como teoria da mente? Bem, a proposta é simplesmente que concebemos a mente humana como uma tabela de ação enormemente complexa, encarnada nos processos neurofisiológicos do cérebro. Como máquinas de Coca, os seres humanos partem de *inputs* na forma de informação sensorial ou perceptual, e os transformam em *outputs* comportamentais. Por exemplo, você recebe pelos seus olhos a informação de que o touro no campo que você está cruzando está prestes a atacar. Essa informação é processada em seu cérebro, resultando em um *output* comportamental, a saber, que você corre para o portão. (A implementação de uma parte diferente da tabela de ação mental poderia, alternativamente, dizer para você ficar onde está se, por exemplo, você teve experiências passadas com touros e sabe como lidar com situações de emergência que os envolva.)

Como já chamei atenção, a tabela de ação que constitui a mente humana seria incrivelmente complexa, mas para o de-

fensor do funcionalismo computacional da máquina de Turing é assim que os seres humanos funcionam internamente. Sem dúvida, a tabela que constitui a mentalidade humana (e a animal) conteria sistemas dentro de outros sistemas, arranjados hierarquicamente, começando com o menor ou mais básico nível com elementos que não podem ser divididos em funções mais simples, e tornando-se gradualmente mais complexo de acordo com a ascensão na hierarquia, até que a estrutura funcional final e completa seja discernida. Tais sistemas complexos são exemplificados pela organização de gráficos de grandes empresas, que detalham as funções e estruturas internas de vários departamentos juntamente com suas relações com outros departamentos, de modo que os objetivos da companhia como um todo possam ser postos em prática. Também máquinas consistem em subsistemas, que permitem que as várias funções da máquina como um todo sejam desempenhadas. No seu livro *Physicalism*<sup>6</sup>, K. V. Wilkes usa o exemplo de uma máquina de lavar para ilustrar essa questão.<sup>6</sup> Máquinas de lavar precisam de sistemas que permitam a entrada de água (*input*), e que despejem a água depois da lavagem (*output*). Elas precisam de um sistema para agitar o tambor onde é feita a lavagem, e de algum meio para aquecer a água. Hoje em dia elas possuem um sistema de programação para permitir a realização de diferentes espécies de rotinas de lavagem. Cada um desses sistemas pode ser dividido, por sua vez, em outros componentes e funções. O tubo de admissão, por exemplo, consistirá em uma válvula operada eletronicamente para abrir e deixar que uma quantidade determinada de água entre na máquina. Por sua vez, a válvula terá vários componentes, até que finalmente os elementos

\*N. de T.: Em português: *Fiscalismo*.

físicos, que não podem ser divididos em unidades funcionais menores, tenham sido alcançados.

No caso dos seres humanos, esses elementos mais simples, pode-se presumir, serão as moléculas (e em última instância os átomos) que constituem as células nervosas que auxiliam a constituir o cérebro e o sistema nervoso central. Em níveis superiores encontraremos estruturas como o hipocampo, que é responsável por armazenar as memórias de longa duração, ou o núcleo geniculado lateral, que serve como ponto de partida para a transferência de impulsos neurais dos nervos ópticos para o lobo occipital do córtex visual, que, sabidamente, lida com a percepção visual. Desse modo, o funcionalismo computacional nos oferece um modelo de processamento de informação da mente que pode explicar nossas habilidades. Afinal, não pode ser por causa de mágica – pó de fadas nas sinapses – que possuímos as capacidades que possuímos para o reconhecimento de faces ou para o uso de linguagem, por exemplo. A teoria computacional da mente nos oferece a melhor chance até o momento para entender por que os seres humanos são do jeito que são. Ou pelo menos é isso que os computacionais dizem.

### Funcionalismo e reducionismo

Assim como o behaviorismo analítico, o funcionalismo é reducionista. Isso quer dizer que ele tenta explicar a natureza dos estados mentais em um vocabulário não-mentalista reduzindo-os a estruturas de *input/output*. Como Sidney Shoemaker observa:

[O] funcionalismo [...] é a doutrina segundo a qual termos mentais, ou psico-

lógicos, são, em princípio, elimináveis de uma certa maneira. Se, para simplificar as coisas, dissermos que nosso vocabulário mental consiste de nomes para estados e relações mentais, [...] a alegação será que esses nomes podem ser tratados como sinônimos de descrições definidas, cada uma das quais pode ser formulada, em princípio, independentemente do uso de qualquer vocabulário mental.<sup>7</sup>

A questão, contudo, é como isso poderia ser feito sem se incorrer em circularidade, isto é, sem se usar os próprios termos que o funcionalismo pretende dispensar com a análise. Como, para tomar o exemplo dado anteriormente, em que alguém vê um touro preparando-se para atacar, e como resultado foge, poderiam os *inputs* e *outputs* ser capturados por termos não-psicológicos? A expressão “ver um touro preparando-se para atacar” explicitamente atribui um estado mental, um estado perceptual, ao indivíduo em pauta. “Fugir” atribui uma ação à pessoa, e desse modo também atribui implicitamente desejos e crenças. A menos que se encontre alguma maneira de descrever esses *inputs* e *outputs* que consiga evitar caracterizá-los mentalmente, a análise funcionalista estará viciada pela importação para seu interior dos próprios termos que buscava eliminar.

Felizmente para o funcionalista, foi encontrado um método por meio do qual o emprego de termos mentais pudesse ser evitado ao caracterizar *inputs* e *outputs*. A técnica emprega um dispositivo que encontramos na página 116, uma “frase de Ramsey”, assim chamada em homenagem a Frank Ramsey, a quem essa invenção é creditada. Para ver como ela funciona, suponha que a pessoa que vê o touro no campo e começa a correr chama-se Ron. Então podemos dizer que Ron tem a crença C de que o touro está se preparando

para atacá-lo devido à sua percepção P de que ele está se movendo rapidamente em sua direção, e isso, somado ao desejo D de não ser gravemente machucado, o leva a desempenhar a ação A de correr para o portão. C, P, D e A são todos termos que direta ou indiretamente atribuem mentalidade a Ron. Como podem o *input* P, os estados internos C e D e o *output* comportamental A ser traduzidos em termos não-mentais?

Podemos começar nos referindo à crença C como um mero "algo", um ou outro estado que resulta da percepção P e que, somado ao desejo D, causa a ação A. Naturalmente, neste estágio, o processo de ramseyficação não está completo, uma vez que a frase ainda contém termos mentais. O próximo passo é substituir outro termo mental, digamos, D, pela expressão "algo mais". Podemos agora dizer que há algo, algum estado que resultou da percepção P, que somado a algo mais, outro estado, causa a ação A. Claramente, o objetivo será no final substituir ambos, P e A, por expressões que evitem caracterizá-los mentalmente. As expressões "algo" e "algo mais" são grosseiras, contudo, e se um grande número de tipos diferentes de estados mentais estão envolvidos, logo nos faltarão expressões adequadas para eles. O que os filósofos fazem é substituir termos mentais como P, C, D e A por variáveis tais como w, x, y, e z, e prefixá-las pelo que se chama de quantificador existencial. O quantificador existencial,  $\exists$ , que é sempre seguido de uma variável, como alguns de vocês talvez já saibam devido ao estudo de lógica, deve ser lido: "Há um w tal que...".

Aplicando essa técnica à frase original e substituindo apenas a percepção do touro, P, por w, obtemos:  $(\exists w)$  (Ron tem w e w causa a crença C, que somada ao desejo D leva à ação A). Substituições ulteriores dos termos mentais C, D e A por

variáveis geram:  $(\exists w)(\exists x)(\exists y)(\exists z)$  (Ron tem w e w causa x, que somado a y causa z). Desse modo as relações causais entre *input*, estados internos e *output* são capturadas sem que se empregue qualquer vocabulário mental, evitando assim a circularidade na análise.

### Pontos fortes do funcionalismo

O funcionalismo possui muitos pontos fortes.

1. Ele evita o erro do behaviorismo de identificar estados mentais com comportamento externo real e potencial, mas também escapa da postulação de misteriosas substâncias anímicas não-físicas caras ao dualista. Adicionalmente, uma vez que funções, por sua própria natureza, são multiplamente realizáveis, ele não se compromete com uma implausível identidade tipotípico mente/cérebro. Desse modo ele também evita a atitude chauvinista de insistir de modo míope que estados mentais só podem existir como estados de cérebros humanos.
2. Pensar em um estado mental em termos de suas causas e efeitos típicos condiz bem com a concepção de senso comum que temos da mente.<sup>8</sup> Uma dor, como vimos, pode ser pensada como uma espécie de estado mental que tipicamente resulta de dano a um tecido e causa comportamento de dor acrescido de um desejo de se livrar da dor.
3. O problema da causação mente/corpo desaparece. Não podemos dizer que a dor, pensada como uma entidade ou estado que existe

independentemente do comportamento de dor, suscita aquele comportamento. Dor é o estado funcional como um todo, caracterizado em termos de *inputs*, *outputs* e relações com outros estados mentais, que também são analisados funcionalmente. O comportamento não é causado pela dor; em vez disso, seria possível dizer que ele é uma parte constitutiva do estado funcional como um todo, que é considerado como estar com dor. Naturalmente, em seres humanos o comportamento de dor será causado por eventos no sistema nervoso central, mas esses eventos tomados isoladamente não podem ser identificados com a dor. Esses eventos ajudam a constituir a corporificação da função por meio da qual certo *input* na forma de dano a um tecido é convertido em certo *output* na forma de contorção e gemido. É o sistema completo de relações entre *inputs* e *outputs* e outros estados funcionais que constitui o estar com dor, e não algum estágio ou aspecto daquele processo.

### **Problemas para o funcionalismo**

A despeito de seus pontos fortes, o funcionalismo enfrenta muitas dificuldades sérias, que analisarei nesta seção.

#### **Liberalismo**

Se o chauvinismo restringe injustamente a mentalidade aos sistemas nervosos humanos, o liberalismo é generoso demais no que toca às atribuições de mentalidade que está preparado para fazer. A essência da posição liberalista é que deve-

se permitir a qualquer sistema funcionalmente equivalente a uma mente humana a posse de estados mentais. Contudo, é possível encontrar muitos exemplos, por vezes bizarros, aos quais é totalmente implausível atribuir mentalidade.

Os dois aspectos cruciais da mente que o funcionalismo parece incapaz de capturar são: (1) a subjetividade e a privacidade dos estados mentais como sensações, seus qualia, em outras palavras; e (2) a intencionalidade ou a sobredade de atitudes proposicionais tais como crenças, desejos e emoções. Dois argumentos – em particular, o assim chamado argumento da mente chinesa, de Ned Block, e o argumento da sala chinesa, de John Searle – foram apresentados para demonstrar a incapacidade do funcionalismo para lidar com qualia e intencionalidade. Isso pode ser resumido diagramaticamente – ver Figura 5.6.

#### **O argumento da mente chinesa: o problema dos qualia ausentes**

O argumento da mente chinesa pode ser explicado em primeira instância por referência ao argumento das cabeças de homúnculos. A falácia do homúnculo foi explicada no Capítulo 2, em seção de mesmo nome. A ideia, como você deve lembrar, é que se você tentar explicar algum processo mental, tal como a visão, postulando um agente, talvez sob a forma de uma alma, no cérebro, que escaneia o *input* visual dos seus olhos, você acaba apresentando uma explicação vazia ou circular, porque você então precisa explicar como esse agente vê. A ameaça é de um regresso infinito de almas dentro de almas.

Ned Block apropria-se da noção de um homúnculo com o fim de articular a seguinte objeção ao funcionalismo.<sup>9</sup> Devemos imaginar um corpo humano ex-

ternamente idêntico ao nosso próprio, mas internamente muito diferente. Seus órgãos sensíveis estão conectados a um banco de luzes na cabeça, e conexões ulteriores com os neurônios de *output* motor são feitas por meio de um conjunto de botões. Dentro da cabeça está um grupo de homenzinhos, ou homúnculos. Um grupo desses homúnculos é chamado de homens-G porque eles produzem operações no que podemos chamar de célula-G, que constitui apenas uma célula da complexa tabela de ação que descreve você. (Lembre-se da simples tabela de ação apresentada anteriormente consistindo em quatro células que caracterizam funcionalmente uma máquina de Coca-Cola). Imagine que um cartão é afixado em um quadro de avisos dentro da cabeça, e que possui um "G" escrito nele. Isso diz aos homens-G para se prepararem. Um *input* luminoso, I<sub>1</sub>, digamos, é acionado fazendo com que um homenzinho pressione o botão de *output* O<sub>200</sub> e altere o cartão, no qual "G" está escrito, para um "M". Em outras palavras, os homenzinhos em sua cabeça executam a tabela de ação que lhe caracteriza funcionalmente, do mesmo modo que um grupo de pessoas poderia ficar dentro da máquina de Coca-Cola e executar a sua tabela de ação. Neste exemplo, "G" e "M" correspondem aos estados internos da máquina que foram caracterizados como E1 e E2 no caso da máquina de Coca-Cola.

Seria plausível supor que esse time de homenzinhos, considerado como um grupo, disfraza de experiências subjeti-

vas? Parece fortemente contra-intuitivo supor que sim. A objeção poderia ser levantada: mas não é verdade que cada um desses homenzinhos possui consciência de modo que talvez, afinal de contas, haja experiência? Mas a objeção não alcança o propósito do exemplo. O que acontece no interior do homenzinho é irrelevante, e a questão é: será que os homenzinhos, simplesmente em virtude de desempenharem a função de alterar células-G para células-M, isto é, de atuarem como os agentes causais que executam o papel funcional que está sendo modelado, desfrutam, portanto, de qualia como um grupo? Novamente, parece até mesmo mais claro que não. Além disso, o que quer que seja colocado no lugar desses homenzinhos continuará não introduzindo qualia. Block em um certo momento supõe que talvez sejamos todos cabeças de pulga e que pequenas pulgas treinadas atuassem no papel de homens-G. Mas nem assim se introduziria qualia. Pode-se argumentar: mas se você tem qualia e descobre-se que era cabeça de pulga, você teria que concordar que cabeças de pulga podem, afinal de contas, possuir qualia. Esse argumento, contudo, já pressupõe o que deveria provar, isto é, que cabeças de pulga possuem qualia. Se eu posso estar seguro de que possuo qualia – e certamente eu posso –, então isso, somado aos experimentos mentais de Block, sugere que simplesmente não se poderia descobrir que eu sou cabeça de pulga ou que existem regimentos de homenzinhos em minha cabeça.

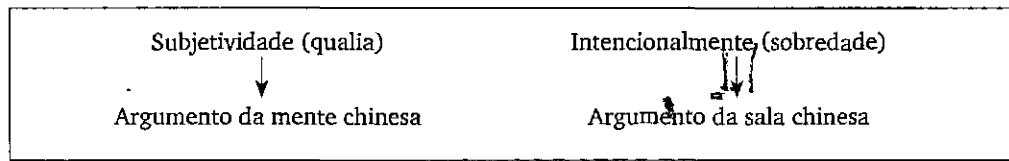


Figura 5.6

Block retoma o argumento que está levantando com o exemplo ainda mais extravagante da mente chinesa, que se tornou uma espécie de pequeno clássico na literatura que trata do funcionalismo. O que agora devemos imaginar é que o governo chinês converteu-se ao funcionalismo e que, com o objetivo de melhorar sua reputação internacional, equipou a totalidade de seus um bilhão de habitantes com transmissores de duas vias que os conectam com outras pessoas e com o corpo artificial descrito no exemplo anterior. O quadro de avisos na cabeça é substituído por satélites que podem ser vistos a partir de qualquer ponto da China. O sistema de um bilhão de pessoas comunicando-se entre si, juntamente com os satélites, constitui um enorme "cérebro" externo artificial.

Tal sistema seria funcionalmente equivalente a um ser humano, como eu ou você, por exemplo. Mas será que ele poderia desse modo desfrutar dos qualia que eu e você desfrutamos? Mais uma vez, a resposta deve ser não. Pode-se levantar a mesma objeção do caso dos robôs cabeça-de-homúnculos: certamente a mente chinesa teria qualia devido aos sinais neurais dos cidadãos chineses que a compõem. Contudo, esses sinais não seriam considerados *inputs* para o sistema, porque o que conta como *input* é estipulado pela pessoa que especifica o que o sistema é. Como afirma Block, os sinais não seriam mais *inputs* do que o seria uma fita inserida por um sabotador entre os contatos do relé no interior de um computador.

O ponto essencial que Block quer estabelecer com todos esses exemplos é que algo poderia ser funcionalmente equivalente ao que se supõe que ocorre no caso das mentes humanas, e mesmo assim não dispor de qualquer vestígio de mentalidade. Não há nada como ser a mente chinesa.

sa, considerada como um todo, e tampouco os indivíduos todos que a compõem possuem qualquer compreensão coletiva do que está se passando. Em conclusão, a mente chinesa não é, portanto, de forma alguma, uma mente.

### Qualia invertidos

Essencialmente o mesmo ponto – que o funcionalismo está deixando algo importante de fora em sua tentativa de explicar a mente – pode também ser estabelecido por referência à possibilidade de qualia invertidos. Você deve lembrar, do Capítulo 1, a metáfora do besouro na caixa de Wittgenstein. Cada um de nós possui o que chamamos de besouro em nossa caixa, mas, uma vez que ninguém pode olhar dentro da caixa de outrem, é possível que as caixas contenham espécies completamente diferentes de coisa, ou até mesmo absolutamente nada. Aplicando-se isso à mente, não será possível que a cor que eu experimento quando vejo uma calêndula seja a cor que você experimenta quando vê uma violeta, e vice-versa?<sup>10</sup> Pode até ser que, indo além, toda sua experiência de cores seja invertida em relação à minha, que quando vejo vermelho você vê verde, quando vejo amarelo você vê azul, e assim por diante. Naturalmente, quando eu e você vemos um objeto vermelho, ambos diremos que ele é vermelho, e quando estivermos decorando uma casa, por exemplo, ambos faremos coisas tais como nos assegurar de que o papel de parede vermelho combine com os tapetes vermelhos, que somente velas vermelhas sejam compradas para as decorações natalinas, e assim por diante. Contudo, suponha que se descubra que, sem que nenhum de nós soubesse, inversores cromáticos foram inseridos em seus olhos quando você nasceu, de modo que desde

então papel de parede, tapetes e velas vermelhas pareceram-lhe como azevinho e árvores de natal pareciam a mim? Os *inputs* podem, além disso, ter originado os mesmos estados internos – digamos, crenças – tanto em seu caso, quanto no meu. Assim ambos seremos levados a crer, e como resultado a julgar, que caixas de correio são vermelhas, mesmo que em seu caso caixas de correio pareçam ter a cor que eu chamaria verde. Também concordaremos que pinheiros de natal são verdes, assim como a grama, embora ambos lhe pareçam da cor que eu chamaria vermelho. Se alguma outra pessoa pedisse a você ou a mim para escolher algo que combine com a cor da grama, ambos poderíamos escolher um ramo do pinheiro de natal, ou pegar uma folha de uma palmeira, muito embora internamente essas coisas pareçam para você da cor que eu chamaria de vermelho.

O que isso demonstra é que podemos ter dois sistemas, você e eu, que são funcionalmente equivalentes em termos de *inputs* e *outputs* (os fótons que entram em nossos olhos, o modo como decoramos a casa para o natal), mas que diferem no que toca a seus estados mentais. Os conteúdos de suas experiências visuais e crenças são diferentes dos meus, mas isso é algo que o funcionalismo não está conseguindo discriminar. Para deixar o que estou dizendo bem claro, o mesmo ponto pode ser expresso de outra maneira, usando o exemplo da dor. É perfeitamente possível imaginar um estado mental que desempenhe o mesmo papel funcional que a dor normalmente desempenha, mas que internamente não se pareça com dor para o sujeito que a tem, e assim não se qualifique como dor, porque o que é essencial para a dor ser dor é simplesmente como ela é sentida. A moral parece clara. Para a vida do ponto de vista mental há mais do que funcionalismo.

### **Qualia ausentes revisitados: o que Fred sabia e Mary não sabia**

Embora estados funcionais não possam, estritamente, ser identificados com estados físicos, eles são corporificados em tais estados, uma vez que propriedades e substâncias não-físicas foram rejeitadas. Nessa medida, os funcionalistas são materialistas. Internamente, não somos nada mais do que cérebros e sistemas nervosos centrais, que recebem um *input* físico por meio dos sentidos na forma de fótons e ondas sonoras, por exemplo, e então convertem esse *input* de várias maneiras em um *output*, na forma de comportamento – movimentos corporais em geral (e imobilidades). O que essa imagem meio tosca sugere é que, uma vez que a investigação científica tenha apresentado os detalhes neurofisiológicos de como os *inputs* são transformados em *outputs*, a tarefa de explicar a natureza da mente terá sido completada. Dizendo de outro modo, de acordo com o materialista ou fiscalista, se pudéssemos conhecer todos os detalhes físicos, todos os fatos físicos conhecidos por meio das ciências como a física, a química, a biologia, etc., sobre o que ocorre dentro de nós quando pensamos, imaginamos, percebemos, acreditamos, sofremos dores e assim por diante, de modo a incluir todos os tipos possíveis de estados e atividades mentais, então saberíamos tudo que há para saber sobre nós. Uma história física completa de nosso funcionamento, cobrindo o que acontece em nossos cérebros e sistemas nervosos centrais, é uma história completa *simpliciter*. Similarmente, se imaginarmos agora esse conhecimento sendo estendido além daquilo que ocorre com, e em, nossos corpos, uma explicação física completa da realidade seria, para um fiscalista dedicado, uma explicação completa, ponto final. Uma vez que conheça-

mos (supondo que isso seja possível) todos os fatos físicos e todas as explicações, os modelos e as teorias que fazem uso apenas de tais fatos (e de nenhum outro, pois não há nenhum outro), então nosso conhecimento seria completo.

Essa alegação foi fortemente desafiada por Frank Jackson em dois artigos muito debatidos e merecidamente famosos: "Epiphenomenal Qualia" e "What Mary Didn't Know"<sup>\*11</sup>. Jackson nos pede para imaginar alguém, Fred, que possui a melhor discriminação cromática já registrada. Ele não apenas pode distinguir cores tão bem quanto qualquer outra pessoa viva, mas também pode fazer discriminações que ninguém mais no mundo pode. Quando todas as demais pessoas olham para tomates maduros, elas veem apenas uma cor. A cor de todos os tomates é exatamente idêntica para elas. Não para Fred, no entanto. Para Fred, a cor dos tomates divide-se em duas classes distintas, vermelho<sub>1</sub> e vermelho<sub>2</sub>, mas, embora por cortesia Fred nos fale de vermelho<sub>1</sub> e vermelho<sub>2</sub>, para ele essas cores não são tons de uma mesma cor em absoluto; são tão diferentes uma da outra quanto a calêndula e a violeta. Quando damos a Fred um monte de tomates, ele pode agrupá-los infalivelmente em dois montes, cada monte composto numericamente das mesmas frutas idênticas, inúmeras vezes, mesmo havendo os tomates sido misturados entre si após cada um dos selecionamentos. Além disso, a investigação do sistema óptico de Fred revela que ele pode separar dois grupos de comprimentos de onda na parte vermelha do espectro, do mesmo modo que nós fazemos ao distinguir o vermelho do verde. Parece necessário admitir que Fred realmente possa ver pelo menos uma cor a mais do que o resto de nós, e isso nos faz ficar pensando como

seria se nós também pudéssemos fazer isso. Caso nossos sistemas nervosos pudessem ser alterados de tal modo que eles se tornassem iguais aos de Fred, então presumivelmente nós, também, seríamos capazes de distinguir tomates assim como Fred, e poderíamos imaginar como seria nossa experiência de ver tomates depois que nossos sistemas nervosos tivessem sido ajustados, comparando-a ao que ela era antes de a mudança ocorrer.

Infelizmente, contudo, não adquiriremos a habilidade de Fred, porque a tecnologia médica não é avançada o suficiente para capacitar cirurgiões a fazer as alterações apropriadas em nossos sistemas nervosos. Assim, ficamos no escuro sobre como seria ver duas cores de tomates completamente diferentes onde antes víamos apenas uma. Pode-se supor que possuímos um conhecimento completo do que acontece fisicamente no cérebro e no sistema nervoso de Fred, assim como um conhecimento completo da estrutura atômica das superfícies de tomates maduros, das propriedades das ondas luminosas e assim por diante, mas nada disso nos possibilitará a experiência de ver as duas cores diferentes que Fred vê. Com relação a Fred, somos como os habitantes da terra dos cegos em relação à pessoa da comunidade que pode enxergar.

O propósito dessa história é este. Embora saibamos tudo que há para saber fisicamente sobre o que acontece no caso de Fred e no nosso, Fred sabe algo que nós não sabemos, ou seja, que tomates, que parecem ter uniformemente a mesma cor para nós, de fato formam dois grupos com duas cores diferentes. Isso implica que o materialismo deixa algo de fora, a saber, como essas cores parecem para um sujeito de uma experiência que pode discriminá-las. Em outras palavras, a imagem materialis-

\*N. de T.: Em português: "Qualia epifenomênico" e "O que Maria não sabia".

ta do mundo não é uma imagem completa, ponto final. É apenas uma imagem física completa. No entanto, a imagem física não é completa. É necessário algo mais; há também o aspecto não-físico, mental, evidenciado pela experiência de Fred de ver duas cores onde vemos apenas uma.

Jackson chega ao mesmo resultado usando o exemplo de Mary, uma cientista brilhante que sempre viveu sozinha em uma sala completamente preto-e-branco, e que foi forçada, por alguma razão, a investigar o mundo por um monitor de TV preto-e-branco. (Suporemos, um tanto artificialmente, que ela não pode ver sequer seu próprio cabelo e corpo, de modo que ela esteja realmente impedida de experienciar qualquer cor exceto a preta e a branca.) Mary especializou-se nos mecanismos do cérebro que lidam com a visão e adquiriu conhecimento de todos os detalhes físicos envolvidos na percepção. Ela sabe quais comprimentos de onda estimulam que partes do cérebro, e como esses acontecimentos, em última instância, levam a pessoa em questão a proferir as palavras "O céu é azul". Mary, podemos supor, pode ver o céu em seu monitor, mas para ela ele parece apenas negro, ou de vários tons de cinza e branco.

Um dia, Mary é libertada de seu quarto, ou ganha uma televisão colorida. Ela olha para o céu e para tomates e exclama: "Então é assim que azul e vermelho parecem para todas aquelas pessoas que eu só pude ver no meu monitor preto-e-branco". Será que Mary aprende alguma novidade, algo que ela não sabia antes de ser libertada? Parece que somos forçados a concluir que sim. Antes, do ponto de vista físico, ela sabia tudo que estava se passando quando as pessoas (e ela mesma) viam cores, mas agora ela aprendeu algo novo, a saber, como é a experiência de cor, que previamente lhe fora negada. Mais uma vez, parece seguir-se que o conhecimento ou informação que ela subsequentemente

adquire não pode ser meramente conhecimento do físico, ou informação sobre ele, porque aquele conhecimento estava completo antes de ela ser libertada.<sup>12</sup>

Você deve agora examinar o Exercício 5.1 antes de prosseguir sua leitura.

### **EXERCÍCIO 5.1**

Sem ler a seguir, tente expor o argumento de Jackson contra o fisicalismo usando duas premissas e uma conclusão. Depois disso leia o que segue.

O argumento de Jackson pode ser apresentado assim:

1. Mary (antes de ser libertada) sabe tudo que há para saber do ponto de vista físico sobre outras pessoas.
  2. Mary (antes de ser libertada) não sabe tudo que há para saber sobre outras pessoas (porque ela aprende algo sobre elas depois de ser libertada).
  3. Portanto, há verdades sobre outras pessoas (e sobre ela mesma) que ficam de fora da história fisicalista.
- Considere agora o Exercício 5.2.

### **EXERCÍCIO 5.2**

Você consegue ver alguma maneira de atacar o argumento de Jackson e mostrar que sua conclusão é falsa? Tente fazer isso em uma discussão com outras pessoas – é bem difícil de conseguir.

Paul Churchland apresentou o que ele chama de uma "versão convenientemente comprimida" do argumento:<sup>13</sup>

1. Mary sabe tudo que há para saber sobre estados cerebrais e suas propriedades.
2. Não é o caso que Mary saiba tudo que há para saber sobre sensações e suas propriedades.

3. Portanto, aplicando a lei de Leibniz (ver Capítulo 2, seção “O argumento da dúvida”), sensações e suas propriedades não são idênticas a estados cerebrais e suas propriedades.

Tendo em mente o argumento de Jackson e a versão dada por Churchland, leia agora o Exercício 5.3.

### **EXERCÍCIO 5.3**

Você concorda que a versão de Churchland capta acuradamente o argumento de Jackson? Você pode sugerir outras críticas a esse argumento? Depois que você tiver pensado sobre isso, leia meus comentários a seguir.

Primeiramente, se o argumento de Jackson se resumisse a não mais do que sugerimos aqui, Churchland o teria demolido. Ele é paralelo ao caso dos dorminhocos (Capítulo 3, seção “Evitando objeções iniciais à teoria da identidade”), que sabem tudo sobre a estrela vespertina e suas propriedades, mas não sabem nada sobre a estrela matutina e suas propriedades (porque estão dormindo quando ela aparece no céu pela manhã), o que supostamente levaria à conclusão de que a estrela vespertina  $\neq$  estrela matutina (porque a estrela vespertina supostamente possui uma propriedade, a saber, ser conhecida pelos dorminhocos, que a estrela matutina não possui). Como vimos, isso não estabelece a não-identidade das estrelas matutina e vespertina mais do que o fato de Lois Lane saber que o homem com quem trabalha é Clark Kent, mas não saber que o homem com quem trabalha é o Super-Homem, prova que Clark Kent não é o Super-Homem.

Mas Jackson argumenta que Churchland não representou acuradamente sua posição. Ele lembra que toda a força de

seu argumento original está em que Mary não sabe tudo sobre estados cerebrais e suas propriedades porque ela não conhece os qualia, as experiências de cor, associados com tais estados. Ela sabe tudo que há para saber do ponto de vista físico sobre eles a partir de sua perspectiva acromática, mas aprende algo novo em acréscimo ao que é fisicamente conhecido quando experimenta cor pela primeira vez. Como Jackson observa, retoricamente, “O que ela sabe de antemão é *ex hypothesi* tudo que há para saber do ponto de vista físico, mas será isso tudo que há para saber? Essa é a questão crucial”.<sup>14</sup>

Como parte de sua crítica, Churchland também indica, acuradamente, que o argumento de Jackson faz uso equivocado da palavra “saber”. O que Mary sabe antes de ser libertada pode ser chamado, seguindo a distinção de Bertrand Russell, de “conhecimento por descrição”, enquanto o conhecimento que ela adquire depois de ser libertada, quando experimenta cor diretamente pela primeira vez, é “conhecimento por familiaridade”.<sup>15</sup> Contudo, isso não é pertinente, e não refuta a tese de Jackson. Como Mary sabe o que ela sabe não vem ao caso, mas o que ela vem a saber, seja como for que ela venha a sabê-lo. Ela efetivamente adquire novo conhecimento depois de ser libertada, algo que ela não sabia antes, e uma vez que o que ela sabia antes era tudo que havia para saber do ponto de vista físico, ela vem a saber algo que a descrição fisicalista deixa de fora.

David Lewis e Laurence Nemirow também argumentaram que ao ser libertada Mary não aprende absolutamente nada novo.<sup>16</sup> Ela meramente adquire uma nova habilidade representacional ou imaginativa. Contudo, embora seja verdade que ela adquire novas habilidades – ela pode, por exemplo, separar perfeitamente vegetais por suas cores, algo em que ela

faria muito mal em sua desbotada existência anterior – e ela agora também pode imágina como sua sala se pareceria coberta com papel de parede carmesim, será que essas novas habilidades são *tudo* que ela adquire?

Jackson argumenta que, se Mary pensasse consigo mesma que desde que foi libertada ela aprendeu sobre as mentes de outras pessoas algo que ela não sabia antes – a saber, como era experimentar cores – e então subsequentemente assistisse a uma aula sobre se podemos realmente saber se mentes além das nossas existem, ela poderia começar a se preocupar com o fato de que talvez, afinal de contas, ela não tivesse adquirido o conhecimento sobre outras pessoas que ela pensou ter adquirido. Suponha que depois disso ela tenha conseguido aquietar seu ceticismo e tenha concluído que ela efetivamente soube algo novo sobre outras pessoas depois de ser libertada, afinal. Então sobre o que, Jackson pergunta, estava ela vacilando – suas habilidades? Certamente não, mas sobre a aquisição de nova informação, de fatos que ela não conhecia antes de ser libertada: “Suas habilidades representacionais eram uma constante de que tinha conhecimento desde o início. Então o que mais a estava torturando, senão a dúvida de ter ou não obtido conhecimento factual dos outros? Não haveria nada com o que se torturar, se habilidade fosse *tudo* que ela adquiriu depois de ser libertada”.<sup>17</sup>

Será que o argumento de Jackson alcançou seu objetivo? Parece-me que sim. Mas aqui eu devo deixar vocês decidirem por si mesmos.

### A sala chinesa e o problema da intencionalidade ausente

Não apenas o funcionalismo é incapaz de capturar os qualia, mas, alega-se,

ele também é incapaz de acomodar outra característica principal de estados mentais: a intencionalidade, a capacidade de atitudes proposicionais serem direcionadas para um conteúdo proposicional, de representarem outros estados de coisas, mesmo aqueles que jamais existiram.

Para estabelecer essa conclusão, John Searle emprega um exemplo que provavelmente tornou-se ainda mais conhecido do que a analogia da mente chinesa de Block: a sala chinesa. O argumento de Searle, como o argumento da seção anterior, visa subverter o funcionalismo metafísico. O ponto que Searle quer estabelecer é que, meramente pela instanciação de um programa ou pela implementação de uma tabela de ação, um computador não pode possuir pensamentos e compreensão genuínos. Um pensamento genuíno possui intencionalidade – ele possui a característica de ser sobre algum estado de coisas diferente dele mesmo. Segundo Searle, um pensamento genuíno possui um significado, um conteúdo semântico. O mesmo se aplica a crenças, emoções, intenções e desejos. Mas as operações desempenhadas por um computador digital não significam nada para ele, e não possuem um conteúdo para ele. As palavras que aparecem no monitor à medida que digito significam algo apenas para você e para mim, mas elas nada significam para aquele processador de texto, porque nele falta um ingrediente essencial que não nos falta: a consciência. Diferentemente de nós, no caso do computador não há ninguém em casa. O computador que mostra em sua tela a mensagem “Olá, Keith” quando eu digito meu nome não está nem me cumprimentando, nem se referindo a mim. Nem poderia uma quantidade qualquer de frases produzidas por um computador, indicando que ele possui certas crenças, temores, intenções, desejos, esperanças e assim por diante significar que

ele literalmente possui aquelas crenças e outros estados intencionais no sentido em que um sujeito humano os possui.

Para demonstrar essa conclusão, Searle nos pede para imaginar que ele está em uma sala da qual não pode sair. Caracteres chineses são introduzidos através de uma fenda em um lado da sala. A tarefa de Searle é correlacionar esses caracteres de entrada com outros que são especificados em um livro de regras. Assim, quando Searle recebe, digamos, um risco-risco, ele procura no livro de regras qual é o caractere com o qual ele deve ser correlacionado, digamos, um rosco-rosco, e passa esse caractere para fora através de uma fenda no outro lado da sala. Dado que Searle-na-sala não possui conhecimento prévio algum de chinês, não há como, meramente por meio desse processo de pareamento, aprender essa língua. Para ele, os caracteres de entrada e de saída são meramente rabiscos sem significado algum, e não palavras. Para os falantes de chinês fora da sala esses caracteres têm significado. Os caracteres introduzidos podem constituir questões em chinês, e os caracteres enviados para fora podem representar respostas a essas questões. O propósito do experimento mental de Searle é mostrar que Searle-na-sala está imitando o que acontece na CPU de um computador, e essa manipulação formal de símbolos não pode levar a qualquer tipo de compreensão dos caracteres que estão sendo manipulados. Meramente em virtude de instanciar um programa dessa maneira, um computador não tem consequentemente como compreender chinês, bem como Searle-na-sala, tampouco.

Não é de muita ajuda imaginar um Searle miniaturizado, homúnculo, dentro do crânio de titânia de um robô com *input*

de caracteres efetuado por câmeras de televisão e os símbolos de *output* servindo para fazer os motores dentro do robô movimentarem seus braços e pernas. Os *inputs* que Searle recebe dos órgãos sensíveis do robô e as instruções que são transferidas para seu aparato motor continuarão não significando nada para Searle-no-robô. Os símbolos que são manipulados por Searle-no-robô não adquirem significado repentinamente pela simples razão de que Searle está agora dentro de algo que se pode mover em meio ao mundo e que pode receber seus *inputs* e transmitir seus *outputs* de maneiras tecnologicamente sofisticadas. Dentro do robô, Searle continua manipulando símbolos formalmente e, por conseguinte, defende Searle, qualquer via à semântica permanece bloqueada.

Nesse exemplo Searle estava disposto a aceitar que, embora programas de computador não possuam uma semântica, no que toca ao computador ou ao manipulador formal de símbolos dentro da sala chinesa, eles ao menos possuem uma sintaxe, isto é, há regras governando as maneiras como símbolos podem ou não ser combinados. O argumento da sala chinesa apareceu já nas *Reith Lectures* de Searle em 1984 e foi publicado subsequentemente como *Minds, Brains and Science*\*.<sup>18</sup> Mas, na época em que Searle escreveu *The Rediscovery of the Mind*\*\*, em 1992, ele deu o passo mais radical de excluir não só a semântica dos sistemas computacionais, mas também a sintaxe.<sup>19</sup> O argumento da sala chinesa mostrou que semântica não é intrínseca à sintaxe. No entanto, agora Searle insiste que a sintaxe, assim como a semântica, não é intrínseca a eventos puramente físicos que ocorrem dentro do computador. Tanto semântica quanto sintaxe são

\* N. de T.: Em português: *Mentes, cérebros e ciência*.

\*\* N. de T.: Em português: *A redescoberta da mente*.

relativas a seres humanos que projetam e usam computadores e fornecem interpretações aos *inputs* e *outputs* dessas máquinas. Como tais, *inputs* e *outputs* não possuem nem características semânticas nem sintáticas. Segue-se então, como indica Searle, que nada é intrinsecamente um computador. Assim, mesmo se pudesse ser provado que certos processos naturais possam ser interpretados como instanciando certas funções, não se seguiria que tais processos poderiam ser descritos como computando aquelas funções. Isso é algo adicionado por um intérprete humano. Para ficar com um exemplo concreto, o *output* de um transformador pode ser caracterizado como o diferencial da corrente alternada ou mutante, mas o transformador não está literalmente usando o cálculo diferencial para converter o *input* em um *output*. Na verdade, nós atribuímos uma interpretação computacional aos *inputs* e *outputs* do transformador: a interpretação existe puramente no que nos diz respeito. Do mesmo modo, os movimentos dos planetas em suas órbitas podem ser computados usando as leis de Kepler dos movimentos planetários, mas os próprios planetas não estão fazendo computações.

Agora, isso que se aplica às entranhas de um computador eletrônico aplica-se igualmente bem ao cérebro, como Searle indica. O cérebro também não pode ser visto como executando funções computacionais e não deveria, ao contrário do que defenderam alguns autores, ser visto como um "mecanismo sintático". "Estados computacionais", diz Searle, "não são descobertos dentro da física, eles são atribuídos à física".<sup>20</sup> Os processos físicos que ocorrem dentro de um sistema tal como um computador ou o cérebro funcionam como causas reais, mas a interpretação funcional ou computacional que atribuímos a tais processos não é intrínseca ao

sistema, ela é antes e meramente algo projetado nele a partir de fora. Consequentemente, a interpretação computacional dos processos físicos não possui existência real, e nem, por conseguinte, poderes causais, uma vez que ela existe apenas nos olhos da pessoa que atribui a interpretação ao sistema.

O ponto que Searle está defendendo pode também ser colocado em termos de sistemas que genuína e efetivamente seguem regras e daqueles que não seguem. Se voltarmos à máquina de Coca-Cola por um momento e imaginarmos que dentro da máquina está de fato uma pessoa que coleta o dinheiro e dá troco e garrafas de Coca-Cola em troca, então é verdade que a pessoa está conscientemente seguindo regras para executar as funções da máquina. A frase "Se £ 1,50 entrarem pela fenda, dê uma Coca e £ 0,50 de troco" possui um significado para a pessoa, e é uma consequência de a pessoa compreender esse significado e decidir seguir a instrução que ele representa, que o troco e a garrafa de Coca sejam entregues. Desse modo, as características semânticas e sintáticas da frase figuram como causas genuínas que ajudam a explicar o comportamento da pessoa.

Mas agora suponha que a pessoa dentro da máquina seja substituída por um mecanismo convencional composto de portinholas e calhas mecânicas ou eletrônicas e assim por diante. Ao entregar uma Coca, esta máquina seguiria as mesmas regras que uma pessoa que estivesse dentro dela? A resposta deve ser não. A máquina aparenta estar seguindo regras, mas ela não pode estar de fato seguindo regras uma vez que, não sendo dotada de consciência, ela não possui compreensão do que está acontecendo quando uma moeda adentra a fenda, e vários processos mecânicos se iniciam, para assegurar que uma garrafa de Coca sairá da máquina.

Podemos falar, e de fato falamos, que a máquina dá troco e entrega Coca para as pessoas, mas esse modo de falar não impõe ações genuínas à máquina e deve ser compreendido meramente como metáfora. A máquina não lhe dá intencionalmente uma Coca, ou intencionalmente deixa de lhe dar troco. Suas operações não podem ser descritas nem como intencionais nem como não-intencionais. A não-devolução de troco seria descrita como devida a um defeito mecânico ou elétrico na operação do mecanismo interno. Isso porque a máquina não tem nenhum ponto de vista, não tem perspectiva de primeira pessoa. Não há nada como ser uma máquina de Coca, diferentemente de humanos ou morcegos.<sup>21</sup> Tudo que está ocorrendo nesse caso são certos processos mecânicos, aos quais nós, como observadores externos, atribuímos um significado.

Um belo exemplo que também ilustra o ponto é fornecido por Anthony Kenny em seu livro *The Metaphysics of Mind*\*.<sup>22</sup> Compare um piano a uma pianola. Quando um pianista toca uma peça musical, isso se deve ao fato de que ele está atribuindo um significado às notas na partitura, isto é, ele está seguindo certas regras para a produção da música. Mas essas notas poderiam ser representadas por furos em um rolo de pianola, e a peça musical poderia ser produzida de maneira puramente mecânica pela operação do rolo posto em movimento. Para a pianola, os buracos não têm o mesmo significado que as notas na partitura têm para o pianista. Os buracos no rolo não são símbolos ou conjuntos de instruções. Mas as notas na partitura são, porque esses símbolos foram estabelecidos por convenções para representar certos sons. Essas convenções foram estabelecidas por seres humanos que podem usar símbolos, para os quais eles possuem um significado, porque seres humanos possuem cons-

ciência e intencionalidade intrínseca. No caso da pianola não-consciente, qualquer compreensão que lhe atribuímos é uma intencionalidade meramente derivativa, ou como se: os buracos no rolo da pianola não são símbolos e não possuem conteúdo semântico, significado ou significância para a pianola. Seguindo as consequências da concepção de Wittgenstein de que os significados dos símbolos não devem ser concebidos como conjuntos de imagens privadas na cabeça, mas como consistindo em usos publicamente conferíveis de símbolos, Kenny acrescenta:

Se meu cérebro fosse tão determinista quanto um computador eletrônico, de modo que todo o seu *output* pudesse ser predito a partir dos *inputs* que recebe, isso ainda não seria suficiente para que alguém estivesse apto a predizer os pensamentos que terei. Pois o que dá significado a qualquer tipo de *output* do meu cérebro – seja ele canalizado por meio de ação, fala ou escrita – é algo completamente externo a ele. O que dá significado a minhas atividades físicas, o que transforma alguns dos sons e gestos que eu produzo em símbolos, é meu poder de participar na atividade social da linguagem – uma atividade que é impossível fora do contexto da cooperação com outros.<sup>23</sup>

### Funcionalismo homuncular

Searle acredita que a única maneira de dotar os processos físicos que acontecem em computadores de uma semântica e de uma sintaxe é colocar um homúnculo – agente consciente e capaz de ter experiência que possui intencionalidade genuína e não derivada – dentro do sistema. Este não é um problema no caso dos PCs que você compra em uma loja, porque cada um deles vem equipado com um homúnculo que

\* N. de T.: Em português: *A metafísica da mente*.

pode interpretar os *inputs* e *outputs* – você! Se ainda assim se quer insistir que os próprios computadores possuem intencionalidade intrínseca, a única possibilidade parece ser colocar um homúnculo dentro do computador. Assim, só falta pensar o cérebro como sendo um computador com um homúnculo posicionado em seu interior, que interpreta os vários processos neurológicos em desenvolvimento, dotando-os de semântica e sintaxe, para retornarmos à completa falácia do homúnculo descrita anteriormente.

O filósofo americano Daniel Dennett propôs uma resposta ousada a essa alegação. Nada há de errado na postulação de homúnculos, contanto que eles sejam mais estúpidos do que a pessoa cujo comportamento inteligente e cuja compreensão os homúnculos sejam invocados a explicar: “Se é possível fazer com que um time de homúnculos relativamente ignorantes, bitolados e cegos produza o comportamento inteligente do todo, isso é um progresso”.<sup>24</sup>

Pode-se pensar em uma pessoa como um conjunto de subsistemas arranjados hierarquicamente, exatamente como a máquina de lavar caracterizada por K. V. Wilkes anteriormente, na seção “Funcionalismo computacional ou da máquina de Turing”. À medida que se desce na hierarquia, os homúnculos tornam-se progressivamente cada vez menos inteligentes até que, na última camada, a homuncularidade e a intencionalidade que ela implica são completamente excluídas do sistema. As genuínas compreensão e capacidade de seguir regras que atribuímos a outras pessoas no nível cotidiano normal devem ser entendidas como o produto do funcionamento do conjunto complexo de subsistemas mais estúpidos. Em outras palavras, afinal de contas, o que acontece nos níveis superiores – o que chamamos de genuínas compreensão e capacidade

de seguir regras – pode ser explicado em termos do comportamento dos elementos de nível inferior em relação aos quais não sentimos inclinação alguma a atribuir intencionalidade. De acordo com essa imagem, realmente não há intencionalidade intrínseca, tal como falam Searle e Kenny, porque ela foi redutivamente decomposta em ações e interações de elementos não-intencionais progressivamente mais simples e básicos. Como observa Searle, “Todos os níveis superiores se reduzem a esse nível inferior; os níveis de cima são todos apenas como-se”.<sup>25</sup>

O problema com uma estratégia redutiva desse tipo é que se fica com a sensação de que ela nada acrescenta à consciência ou à compreensão do ser humano que está sendo decomposto em subsistemas progressivamente mais estúpidos. Em resposta a isso, Dennett sugeriu a analogia de uma grande organização, em que todos os vários trabalhadores e departamentos executam suas funções inferiores, enquanto um secretário de imprensa representa o que está acontecendo no mundo externo. Infelizmente, a analogia falha, porque organizações como um todo não são conscientes ou autoconscientes. Além disso, a importância de um secretário de imprensa – que poderia ser chamado de Homúnculo Mestre, combinando o trabalho de todos os homúnculos mais estúpidos das posições inferiores e representando-o ao mundo externo – simplesmente levanta a questão de como, por sua vez, o próprio Homúnculo Mestre está consciente. A ideia de um tal homúnculo, é interessante notar, não é nova, e pode ser encontrada no trabalho de William James, escrito há mais de cem anos: “Cada célula do cérebro possui sua própria consciência individual, da qual nenhuma outra célula possui conhecimento... Há, contudo, em meio às células, uma central ou pontifícia, à qual nossa consciência está

atrelada".<sup>26</sup> O próprio James, no entanto, rapidamente colocou essa teoria de lado, não apenas porque ela leva a um regresso infinito de homúnculos, mas também porque não há evidência de que exista qualquer coisa como uma célula pontifícia no cérebro: "Não há célula ou grupo de células no cérebro com tal preeminência anatômica ou funcional que lhe faça parecer como a pedra angular ou o centro de gravidade de todo o sistema".<sup>27</sup>

### **Questões para reflexão**

1. Searle está realmente defendendo que, em última instância, organizações de matéria, não importando o quão complexas, nunca serão capazes de pensamento e consciência?
2. Um fisicalista pode explicar completamente o que está acontecendo quando alguém é expulso por cometer uma falta no jogo de futebol?
3. Será que uma moeda de £ 0,50 é apenas um objeto físico?
4. Será que qualia realmente existem?
5. Se uma função não é mental nem física, o que é então?

### **Sugestões para leituras aprofundadas**

- LYCAN W. *Mind and Cognition*. Oxford, Blackwell, 1990 constitui excelente coleção de artigos sobre funcionalismo e assuntos relacionados. A coleção inclui as tentativas de Daniel Dennett, contra Frank

Jackson, de se livrar dos qualia, de "quineá-los" da existência – ver "Quining Qualia", p. 519-47.

- DENNETT D. *Consciousness Explained*. Harmondsworth, Penguin, 1991 tem concepções que tendem para o fisicalismo e para o eliminativismo.
- CRANE, T. *The Mechanical Mind*. Harmondsworth, Penguin, 1995 é uma das introduções mais claras que conheço ao funcionalismo da máquina de Turing e assuntos relacionados.
- MITCHELL, D. B. e JACKSON, F. *Philosophy of Mind and Cognition*. Oxford, Blackwell, 1996 é excelentemente claro e abrangente.
- SEARLE, J. *The Rediscovery of Mind*. Cambridge, MA, MIT Press, 1992 é bastante legível e polêmico.
- SEARLE, J. *The Mystery of Consciousness*. London, Granta Books, 1997. Este livro é hilariante – mas sério – retratando trocas muito espirituosas entre Searle, Dennett e David Chalmers. Uma leitura muito acessível e que vale a pena ser feita.
- DAVIES, M. e HUMPHREYS, G.W., eds. *Consciousness*. Oxford, Blackwell, 1993 é uma coleção útil.
- MADELL, G. *Mind and Materialism*. Edinburgh, Edinburgh University Press, 1988 apresenta boas discussões e ataques a várias teorias materialistas da mente.
- ROSENTHAL, D., ed. *The Nature of Mind*. Oxford, Oxford University Press, 1991 apresenta uma coleção abrangente de artigos agrupados em seções diferentes, com introduções claras às seções feitas pelo próprio Rosenthal.

## NOTAS

1. Discussões excelentes da natureza de pessoas podem ser encontradas nos seguintes textos: DENNETT, Daniel "Condition of Personhood", in: *Brainstorms*. Brighton, Haerwester Press Ltd., 1981, p. 267-85; FRANKFURT, Harry, Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, 68, 14 de Janeiro de 1971; STRAWSON P. F. "Freedom an Resentment", in: *Freedom and Resentment and other essays*. London, Methuen and Co, 1973; O'HEAR, Anthony, ed. *Minds and Persons*. Cambridge, CUP, 2003; SCRUTON, Roger, *Animal rights and wrongs*. London, Damos, 1996.
2. Compare com a discussão de Donald Davidson em *Subjective, Intersubjective, Objective*. Oxford, Clarendon Press, 2001, p. 97-100.
3. STRAWSON, P. F. *The Bounds of Sense*. London, Routledge, 1966, p. 273.
4. CRANE, T. *The Mechanical Mind*. Harmondsworth, Penguin, 1995, p. 93-9.
5. BLOCK, N. "Troubles with Functionalism", in: BLOCK, N., ed. *Readings in the Philosophy of Psychology*. vol. 1 London, Methuen, 1980.
6. WILKES, K. V. *Physicalism*. London, Routledge and Kegan Paul, 1978.
7. SHOEMAKER S. "Functionalism and Qualia", in: BLOCK, N., ed. *Readings in the Philosophy of Psychology*. vol. 1 London, Methuen, 1980, p. 261.
8. BRADDON, D. M. e JACKSON, F. *Philosophy of Mind and Cognition*. Oxford, Blackwell, 1996. Os autores optam assim por chamar o funcionalismo metafísico de funcionalismo de senso comum.
9. BLOCK, N. "Troubles with Functionalism", p. 276.
10. Essa possibilidade foi debatida primeiro por John Locke (1632-1704) em seu *Ensaio Sobre o Entendimento Humano*, Livro II, xxxii.15: é possível que "as diferentes estruturas de nossos órgãos fossem de tal modo ordenadas que o mesmo objeto devesse produzir nas mentes de vários homens diferentes ideias ao mesmo tempo, tal como se a ideia de uma violeta produzida na mente de um homem por seus olhos fosse a mesma que uma calêndula produzisse na mente de outro homem, e vice-versa".
11. JACKSON, F. Epiphenomenal Qualia, in: LYCAN, W., ed., *Mind and Cognition*. Oxford, Blackwell, 1990, p. 469-77; e "What Mary Didn't Know", in: ROSENTHAL, D., ed. *The Nature of Mind*. Oxford, OUP, 1991, p. 392-4.
12. O caso imaginário de Mary recebe suporte de um dramático incidente da vida real. SACKS, O. *An Anthropologist on Mars*. London, Picador, 1995. O autor descreve o caso do Sr. I, que fora um artista bem-sucedido, mas depois de um acidente de carro, repentina e permanentemente, perdeu toda sua visão de cores. Em uma carta ao Dr. Sacks, o Sr. I escreve:
 

"Minha visão era tal que tudo me aparecia como se eu estivesse assistindo a uma televisão em preto-e-branco. Dentro de alguns dias, eu podia distinguir letras e minha visão tornou-se como a de uma águia – eu posso ver um verme serpenteando a uma quadra de distância. A precisão do foco é incrível. MAS – ESTOU ABSOLUTAMENTE CEGO PARA CORES. Eu visitei oftalmologistas que não sabem nada sobre esse assunto de cegueira de cores. Eu visitei neurologistas sem nenhum resultado. Sob hipnose, eu continuo não podendo distinguir cores. Andei me envolvendo em todo tipo de testes. Qualquer um que você imaginar. Meu cão marrom é cinza escuro. Suco de tomate é preto. TV a cores é uma sopa confusa", p. 288.

O desafortunado Sr. I continuou descrevendo o mundo cinzento que ele passou a habitar como não sendo apenas um mundo onde as cores estão ausentes, mas como um onde o que ele via tinha "uma aparência desagradável, "suja", os brancos brilhando, mas descoloridos e encardidos, os negros cavernosos – tudo errado, não-natural, manchado e impuro" (*ibid.*, p. 5).

O caso do Sr. I é o contrário do de Mary. Ela vem a saber como é experimentar cores, enquanto ele perde a habilidade de saber isso. Mas a questão filosófica continua de pé: se a experiência da espécie humana toda tivesse sido como a do mundo visual do Sr. I desde o começo, poderíamos nós saber, puramente baseados no conhecimento de nossa fisiologia, como é ver o vermelho flamejante do pôr-do-sol, o azul-celeste iridescente de uma borboleta *Morpho*, ou o verde fresco das folhas de um *Aesculus* na primavera?
13. CHURCHLAND, P. Reduction, Qualia and the Direct Introspection of Brain States. *The Journal of Philosophy* 82, Janeiro de 1985, p. 8-28.
14. JACKSON, F. "What Mary Didn't Know", p. 393.
15. RUSSELL, B. *The Problems of Philosophy*. Oxford, OUP, 1959, p. 25-32 apresenta a distinção entre "conhecimento por descrição" e "conhecimento por familiaridade".

16. LEWIS, D. "Mad Pain and Martian Pain". [Pós-escrito], in: *Philosophical Papers*, vol. 1. Oxford, OUP 1983, p. 130-2, e "What Experience Teaches", in: LYCAN, W. *Mind and Cognition*, p. 499-519; NE-MIROW, L. Resenha de NAGEL, T. *Mortal Questions*, *Philosophical Review*, 1980, p. 475-6, e "Physicalism and the Cognitive Role of Acquaintance", in: LYCAN, W. *Mind and Cognition*, p. 490-9.
17. JACKSON, F. "What Mary Didn't Know", p. 394.
18. SEARLE, J. *Minds, Brains and Science*. Harmondsworth, Penguin, 1984.
19. SEARLE, J. *The Rediscovery of Mind*. Cambridge, MA, MIT Press, 1992, p. 210.
20. Ibid., p. 213.
21. A expressão "algo como ser nós" ["*something it is like to be us*"] é tomada do famoso ensaio de Thomas Nagel "What is like to be a bat?" ["*Como é ser um morcego?*"]. Nagel defende que nenhuma quantidade de informação obtida sobre uma pessoa de uma perspectiva objetiva, de terceira pessoa, pode nos dizer como é ser aquela pessoa. Como é ser uma pessoa é algo que pode ser conhecido apenas de uma perspectiva subjetiva, de primeira pessoa. Similarmente, como é ser um morcego só pode ser conhecido do ponto de vista subjetivo do morcego. Nós somos tão diferentes de morcegos, e morcegos são tão diferentes de nós, no que toca aos modos e capacidades sensoriais e intelectuais, que jamais compreenderemos completamente a natureza de uma experiência de morcego como ela parece a um morcego, e morcegos, do mesmo modo, jamais saberão propriamente como é ser nós, de nossa perspectiva. O artigo de Nagel, que vale a pena ser lido, pode ser encontrado, entre outros lugares, em seu *Mortal Questions*. Cambridge, CUP, 1979, p. 165-80.
22. KENNY, A. *The Metaphysics of Mind*. Oxford, Clarendon Press, 1989, p. 152.
23. Ibid., p. 153.
24. DENNETT, D. *Brainstorms*. Brighton, Harvester Press, 1978, p. 123.
25. SEARLE, J. *The Rediscovery of Mind*, p 213.
26. JAMES, W. *The Principles of Psychology*, 3 volumes. Cambridge, MA, Harvester University Press, 1976; publicado originalmente em 1890, p. 179, apud: FLANAGAN, O. *The Science of the Mind*. Cambridge, MA, MIT Press, 1984, p. 42.
27. Ibid.

# 6

## LEVANDO A CONSCIÊNCIA A SÉRIO: MONISMO NÃO-REDUTIVO

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender o que pretende o monismo não-redutivo.
- Ser capaz de expor a doutrina da superveniência e dar alguma explicação sobre os aspectos da superveniência.
- Ser capaz de explicar a "solução" de Searle para o problema mente/corpo e propor críticas a ela.
- Ser capaz de apreciar alguns dos problemas encarados pelo monismo não-redutivo.
- Entender por que McGinn pensa que nós nunca seremos capazes de explicar como o cérebro dá origem à consciência de criticar essa posição.
- Compreender uma abordagem de base humeana ao problema da consciência.

### Introdução

Até aqui, examinamos e encontramos insuficiências nas seguintes teorias sobre a mente: a teoria da identidade, tanto na versão tipo-tipo quanto na espécime-espécime; o eliminativismo; o behaviorismo analítico; e o funcionalismo. O principal problema com todas essas teorias, que são ou declaradamente materialistas ou materialistas na inspiração, é que elas não conseguem levar a consciência a sério, seja por não abrir espaço para ela ou por explicitamente negar a sua existência.

O dualismo de substância ao menos admitiu uma realidade plena aos estados mentais concebidos como não-físicos em sua essência. Mas os problemas com o dualismo são numerosos, como temos visto, e entre eles há dois que são particularmente intratáveis. O primeiro é a questão de como almas incorpóreas devem ser individuadas e identificadas ao longo do tempo: a indisponibilidade de tais critérios, por si só, parece desferir um golpe mortal no dualismo de substância. O segundo problema diz respeito a como se pode supor que as almas desfrutem de uma relação

causal com o corpo se as duas entidades, supostamente relacionadas causalmente, são tão completamente diferentes uma da outra em suas naturezas. Explicar como uma entidade não-extensiva poderia interagir causalmente com uma extensiva provou ser uma dificuldade insuperável para Descartes, e assim permanece até nossos dias.

### Dualismo de propriedades e monismo não-redutivo

A que teoria da relação mente/corpo, então, podemos nos voltar agora? O que parece ser requerido é uma teoria que possa trilhar um caminho intermediário entre o materialismo radical e o dualismo forte, uma teoria que, por um lado, não busque negar o fato da existência da mentalidade reduzindo estados mentais a estados puramente físicos, mas, por outro lado, não transforme os possuidores de estados mentais em fantasmas cartesianos incorpóreos, incapazes de afetar o mundo.

Um candidato que parece atender à necessidade é apresentado pelo dualismo de propriedades. O dualismo de propriedades, que foi descrito brevemente no Capítulo 2, seção "Dualismo substancial", exemplifica o monismo não-redutivo. Ele é não-redutivo porque não insiste que as propriedades mentais não sejam nada além de propriedades físicas. Ao contrário, ele está disposto a admitir que propriedades mentais sejam diferentes, em espécie, de propriedades físicas, e não-redutíveis, ontologicamente, a elas. São grupos e séries dessas propriedades mentais que constituem nossas vidas psicológicas. Por outro lado, o dualismo de propriedades não precisa de um dualismo de substâncias – substâncias materiais e substâncias imateriais. Existem apenas substâncias físicas e eventos mentais, e

por essa razão é uma forma de monismo. Mas essas substâncias físicas e esses eventos possuem duas espécies muito diferentes de propriedades: propriedades físicas e, além destas, propriedades mentais, não-físicas, por isso o título da teoria. Na versão do dualismo de propriedades que estou apresentando para inspeção, essas duas espécies de propriedade não coexistem por acaso, produzindo apenas uma relação acidental de uma propriedade com a outra. Em vez disso, há uma dependência unilateral do mental sobre o físico, uma relação que veio a ser conhecida na literatura que envolve este tópico como relação de superveniência. Uma caracterização completa da superveniência é apresentada na próxima seção.

### A doutrina da superveniência

Na recente filosofia da mente, Donald Davidson é um bom representante do tipo de posição monista não-redutiva que há pouco esbocei. Como vimos no Capítulo 3, seção "Racionalidade e normatividade", Davidson negou que a redução completa do mental ao físico fosse possível. Leis de ponte conectando eventos sob suas descrições mentais com aqueles eventos descritos fisicamente foram descartados, e Davidson abraçou uma versão da teoria da identidade espécime-espécime. Para Davidson, eventos devem ser pensados como particulares concretos, não-reproduzíveis, estando no mesmo nível que substâncias lógicas. Visto que ele rejeitou o dualismo de substâncias, não há eventos puramente mentais, mas apenas eventos puramente físicos. Ele sustenta, porém, que esses eventos físicos não possuem apenas propriedades físicas, mas também propriedades mentais irredutíveis. Dessa maneira ele veio a esposar o monismo não-redutivo.

Mas simplesmente afirmar que espécies de eventos possuem tanto propriedades mentais irredutíveis quanto propriedades físicas foi considerado profundamente não atraente, pois deixa a relação das propriedades mentais com as propriedades físicas sem explicação (veja Capítulo 3, seção “A identidade espécime-espécime é o prospecto de redução”). Nós devemos supostamente aceitar, como um fato irracional, que essas duas espécies de propriedades coexistem como aspectos dos eventos, um não explicável em termos do outro, como um objeto que poderia simplesmente acontecer de ser tanto vermelho quanto quadrado, e esses aspectos não possuem qualquer relação intrínseca ou explicativa um com o outro?

Para sair dessa dificuldade e colocar as propriedades mentais em uma relação mais próxima com as propriedades físicas, sem reduzi-las a propriedades físicas ou com elas identificá-las, Davidson propõe a seguinte relação entre as duas espécies de propriedades. Propriedades mentais, disse ele, são supervenientes a propriedades físicas. Traduzindo literalmente, esta palavra incomum e francamente obscura significa “levar ao topo de”. A ideia essencial é a de que um fenômeno, o fenômeno superveniente, surge “em cima de” um fenômeno mais básico ao qual deve seu caráter e existência, o fenômeno de base, subveniente.

A explicação do próprio Davidson sobre a superveniência apresenta-se assim:

Apesar do fato de que a posição que eu descrevo nega a existência de leis psicofísicas, ela é consistente com a visão de que as características mentais são em algum sentido dependentes, ou supervenientes, de características físicas. Tal superveniência pode ser tomada como significando que não podem existir dois eventos semelhantes em todos os aspectos físicos, mas que difiram em alguns aspectos mentais, ou que um objeto não possa ser alterado

em algum aspecto mental sem alterar-se em algum aspecto físico.<sup>1</sup>

Essa caracterização da superveniência é extremamente concisa. Realmente, como Jaegwon Kim observou, no contexto da obra de Davidson isso soa quase como uma explicação posterior, uma observação descartável.<sup>2</sup> Entretanto, à luz da grande quantidade de tinta que tem sido derramada desde então a respeito da superveniência, essa caracterização ainda consegue capturar o que é essencial à relação de superveniência. A caracterização de Davidson da superveniência, juntamente com as observações formuladas a partir do contexto em que ela ocorre, parece possuir três elementos.

**Irredutibilidade:** Os fenômenos supervenientes, ou os fatos, não são reduzíveis analiticamente, por definição, ou ontologicamente, a fatos ou fenômenos subvenientes. Os fenômenos supervenientes existem além dos fenômenos subvenientes de base. Certamente, no contexto da teoria espécime-espécime, ainda que não em alguma outra caracterização mais ampla da superveniência, as propriedades mentais não podem ser reduzidas por definição a propriedades físicas, e também não podem ser ontologicamente reduzidas, ao contrário do que a teoria tipo-tipo sustenta.<sup>3</sup> A distinção explícita, elaborada por Davidson, entre propriedades físicas e mentais, na passagem anterior, apoia esta caracterização da relação de superveniência.

**Covariância:**<sup>2</sup> Os fenômenos supervenientes são determinados por mudanças na base subveniente subjacente, e com elas covariam. Isso significa que podem existir mudanças nos fenômenos supervenientes se, e apenas se, existirem mudanças correspondentes nos fenômenos subvenientes de base. O contrário não se

segue: uma mudança na base subveniente não significa, necessariamente, uma mudança nos aspectos supervenientes. No caso da relação mente/corpo, a razão pela qual o físico pode mudar sem que o mental mude é fornecida pela tese da múltipla capacidade de realização que leva à identidade espécime-espécime – um e o mesmo tipo de estado mental é realizável por meio de uma variedade de tipos diferentes de combinações físicas (ver Capítulo 3). Outra maneira de explicar o que uma covariação significa é dizer que dois indivíduos não podem diferir em suas propriedades supervenientes a menos que haja alguma diferença em seus aspectos subvenientes. Em outras palavras, se dois indivíduos parecem ser exatamente iguais em seus aspectos subvenientes, eles devem ser indiscerníveis a respeito de suas propriedades supervenientes. Uma duplicata física de você, que é exata, molécula por molécula, deve assemelhar-se a você mentalmente.

**Dependência:** Os fenômenos supervenientes emergem, e são dependentes em sua existência, de fenômenos subvenientes de base. Eles também devem seu caráter particular a esses fenômenos de base. Essa dependência é assimétrica. No contexto da teoria da identidade espécime-espécime, a existência do mental e de

seus aspectos é determinada pelo físico. Nessa teoria, a prioridade conferida ao físico deixa claro que ela é uma forma fraca de fisicalismo. Ao físico é dada uma importância primordial, mas sem ir tão longe a ponto de reduzir o mental ao físico ou eliminar o mental em favor dele. Por isso essa aplicação particular da superveniência permite um monismo não-redutivo ou um dualismo de propriedades. Há duas espécies distintas de propriedades, físicas e mentais, mas no final o mental está ali apenas por causa do físico (compare minhas observações sobre a teoria evolutiva, Capítulo 3, seção “A simetria de afirmações de identidade”), e não pode ser alterado independentemente dele.

A superveniência, como descrita aqui, pode ser resumida em um diagrama – veja Figura 6.1. Porém, isso provavelmente soa muito abstrato, então um exemplo pode ajudar.

Os aspectos da superveniência, como definimos anteriormente, são todos satisfeitos pela famosa “alegoria da caverna” de Platão, na *República*. Nessa alegoria, Platão representa os não-filósofos, que estão no mais baixo estágio de cognição – a imaginação – por meio de prisioneiros acorrentados em uma caverna desde o nascimento, de tal forma que eles apenas podem ver o fundo da caverna à sua frente. Às costas desses prisioneiros há

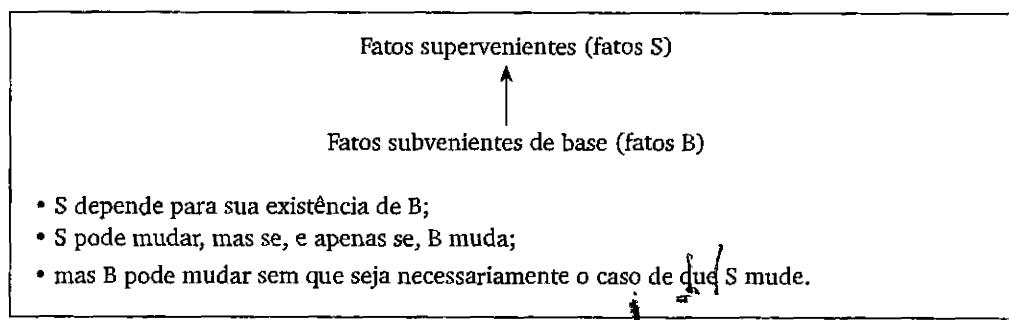


Figura 6.1 Superveniência.

uma parede de cortina, atrás da qual há uma fogueira. Entre a fogueira e a parede de cortina, homens estão caminhando e transportando todas as espécies de objetos sobre suas cabeças. As sombras desses objetos são projetadas pela fogueira no fundo da caverna, que está diretamente na frente dos prisioneiros. As sombras não são substâncias lógicas genuínas, capazes de existir por si mesmas, mas dependem para sua existência de objetos físicos e da fogueira, e a respeito da existência destes dois elementos os prisioneiros são ignorantes. Os prisioneiros jogam um jogo tentando predizer a ordem de aparecimento das sombras e, se alguma regularidade nas aparições é descoberta, eles poderiam muito bem concluir que as sombras anteriores causavam a ocorrência das últimas sombras. Isso seria um equívoco, assim como, muito claramente, seria supor que a existência das sombras e de suas sequências é determinada inteiramente por objetos físicos (e a fogueira) que projetam as sombras. As sombras são distintas dos objetos projetados nas sombras e irredutíveis a eles. Qualquer tentativa de redução seria incoerente, pois isso implicaria vincular a identificação de uma sombra com o objeto do qual ela é uma sombra.

A alegoria de Platão também ilustra a covariância. As sombras podem mudar se, e somente se, os objetos físicos e a fogueira projetando as sombras são alterados de alguma maneira – isto é, se os objetos se movem ou se a fogueira se move. Além disso, dois objetos exatamente similares fisicamente, conduzidos pela fogueira atrás deles de maneira idêntica, e na mesma distância deles, devem projetar a mesma espécie de sombra. (Estritamente falando, os objetos deveriam estar sucessivamente na mesma posição, uma vez que dois objetos não podem ocupar o mesmo espaço ao mesmo tempo.)

### Por que a superveniência prevalece?

A natureza da relação superveniente tem sido detalhada formalmente. Mas por que um par de fenômenos deve ser relacionado dessa maneira? Se, e quando, a superveniência prevalece, qual é a razão? Um dos principais expoentes da superveniência, Jaegwon Kim, cujo trabalho eu referi anteriormente, tem apontado que o mero recurso à relação de superveniência não explica, por si mesmo, por que a superveniência se sustenta. A superveniência, defende Kim, não é uma teoria da relação mente/corpo que corresponda inteiramente às outras principais teorias que examinamos, mas deixa aberta a razão precisa pela qual o mental sobrevém ao físico.

Provavelmente a resposta mais direta a essa questão é a de que o fenômeno superveniente é causalmente dependente do fenômeno subveniente, mas a base subveniente tem de ser concebida como uma causa de sustentação que precisa operar continuamente a fim de manter o fenômeno superveniente existindo. (Uma analogia possível aqui é aquela com uma fonte de água que precisa fluir continuamente para sustentar uma bola de pingue-pongue no ar.) Certamente, no caso das sombras sobre a parede na caverna de Platão, é claro que a relação entre a fonte de luz e os objetos, por um lado, e as sombras, por outro, é de dependência causal. Retire a luz e os objetos e você retirará as sombras. É tão simples quanto isso.

Porém, se a superveniência for enfraquecida para excluir a condição de dependência explicada anteriormente, a fim de que apenas os aspectos da irredutibilidade e da covariância permaneçam, então o paralelismo psicofísico de Leibniz (veja Capítulo 2, seção “Outras tentativas de explicar a relação mente/corpo”) ilustra

esta forma modificada de superveniência. O mental e o físico formam dois sistemas completos e incompatíveis; consequentemente, um não pode ser identificado com o outro. Mudanças mentais covariam com mudanças físicas (ainda que, igualmente, mudanças físicas covariem com mudanças mentais). Mas, em razão da própria natureza do paralelismo, nenhum dos domínios, seja o mental ou o físico, pode ter o mais leve efeito um sobre outro, eles meramente ocorrem como ocorrem, pertencendo à ação inicial de Deus a sincronização dos relógios mental e físico.

A filosofia da mente não é a única área em que a noção de superveniência tem sido empregada. Em seu livro *The Language of Morals*<sup>\*</sup>, R. M. Hare apelou para a noção de superveniência quando sustentou que avaliações morais sobrevêm a fatos não-morais subjacentes.<sup>4</sup> Hare concorda com Hume que nenhuma quantidade de fatos não-morais pode alguma vez implicar logicamente a existência de quaisquer avaliações morais; consequentemente, não existe dependência lógica do avaliativo sobre o factual. Mas o factual determina o moral no sentido de que qualquer mudança em uma avaliação moral de uma situação requer alguma mudança nos fatos que compreendem aquela situação. Hare fornece o exemplo de um julgamento moral sobre a vida de São Francisco para ilustrar o ponto:

Suponha que nós digamos "São Francisco foi um bom homem". É logicamente impossível dizer isso e sustentar ao mesmo tempo que poderia ter existido outro homem colocado exatamente nas mesmas circunstâncias que as de São Francisco, e que nelas se comportasse exatamente da mesma maneira, mas que diferisse de São Francisco apenas a respeito de uma coisa: que ele não fosse um bom homem.<sup>5</sup>

Avaliações morais não são redutíveis, por definição ou análise, a fatos não-morais; supor de outra maneira seria cometer o erro do naturalismo ético, que tenta reduzir julgamentos de valor a afirmações factuais. Porém, não pode ser dito que fatos não-morais causam avaliações morais. A razão pela qual a relação se sustenta nesse caso, como Burwood, Gilbert e Lennon deixam claro, é que "A relação entre propriedades morais e naturais é construída em nosso vocabulário moral. Alguém não dominou a prática de dar justificações morais se ele não aceitou as exigências da relação de superveniência."<sup>6</sup> Em outras palavras, uma pessoa não entende o que significa justificar uma ação moralmente se ela não percebe que uma mudança em um julgamento requer a menção de uma mudança nos fatos não-morais moralmente relevantes, e que precisamente os mesmos julgamentos morais devem ser aplicados nestes casos, como aquele de São Francisco referido, em que todos os fatos não-morais que fundamentam o julgamento são relevantemente parecidos.

De estilo semelhante, os aspectos das obras de arte têm sido ditos ser supervenientes aos vários aspectos físicos que ajudam a abranger obras de arte. É por isso que não pode haver uma diferença estética entre, digamos, duas pinturas, a menos que haja alguma diferença nas propriedades subjacentes às quais os aspectos estéticos das pinturas sobrevêm. O que exatamente constitui os aspectos supervenientes de base de uma obra de arte é uma questão complexa, que não podemos levar adiante aqui. (Parece que isso inclui propriedades históricas, não observáveis, tais como vir da mão de uma pessoa em vez da de outra, assim como aspectos diretamente observáveis.) Mas a menos que haja *alguma* diferença entre quaisquer

\*N. de T.: Em português: *A linguagem da moral*.

aspectos de base que estejam em questão no caso de duas pinturas, os aspectos estéticos das pinturas, e seus valores como obras de arte, precisam ser os mesmos.

Parece-me que até que sejam esclarecidas as razões pelas quais a relação de superveniência se sustenta em espécies particulares de casos, isso deve permanecer com um valor limitado na explicação da relação entre o mental e o físico. Isso é, evidentemente, aceitando que isso se mantém entre a mente e o corpo, algo que Descartes – que acreditou que o corpo físico e a alma imaterial podem ir por seus próprios caminhos em uma gloriosa independência uma da outra – teria terminantemente negado. Em contradição ao cartesianismo, porém, não parece desarrazoado sustentar que o físico seja ontologicamente mais básico, e que enquanto possam existir fenômenos físicos sem fenômenos mentais, a teoria evolucionária (veja Capítulo 3, seção “A simetria de afirmações de identidade”) dá todas as razões para acreditar que, na realidade, não existiriam mentes sem arranjos complexos de matéria para dar surgimento a elas. Podem existir mundos possíveis em que existem estados mentais que não sejam produzidos por processos físicos subjacentes. Essa possibilidade, ao que parece, não pode ser logicamente excluída. No mundo real, porém, é verdade que, tanto quanto podemos dizer, o mental não estaria lá se não fossem os tipos certos de traços físicos para dar surgimento a ele.

### Searle e o problema mente/corpo

De forma tipicamente contundente, John Searle toma o ponto nevrálgico da relação entre a mente e o corpo com as duas mãos. De acordo com ele, a solução do problema é extremamente simples e

tem estado à disposição por muitos anos. Aqui está a solução: os cérebros causam as mentes. Mais exatamente, diz Searle, os fenômenos causais são provocados por processos neurofisiológicos no cérebro e são, eles mesmos, aspectos do cérebro. Isso é o que Searle chama de “naturalismo biológico”: “eventos e processos mentais são tanto parte de nossa história biológica natural quanto a digestão, a mitose, a meiose ou a secreção de enzimas”<sup>7</sup>.

Fazendo uma breve digressão, o ponto da palavra naturalismo é assinalar que eventos mentais fazem parte integral do mundo natural, e não estão, em sentido algum, além ou fora dele – isto é, não não-naturais em sentido algum, e certamente não sobrenaturais. Sua existência e caráter são, em princípio, completamente explicáveis em termos das ocorrências mundanas, naturais. O naturalismo, portanto, não está separado pela distância de um milhão de milhas do projeto do fisicalismo, do objetivo de explicar tudo, no final das contas, em termos dos conceitos e teorias extraídas das ciências naturais (veja Capítulo 1, seção “Intencionalidade”).

Porém, retornando à exposição dos pontos de vista de Searle, ele passa a apresentar com mais detalhes o que ele quer dizer quando chama os fenômenos mentais de aspectos do cérebro. Ele não está contente em apenas dizer que esses fenômenos são propriedades do cérebro, mas afirma, em vez disso, que os dois são causados por microestruturas no cérebro, especificamente a estrutura neuronal, e são por elas realizados. Em uma tentativa de tornar o significado dessa alegação tão clara quanto possível, Searle usa uma analogia. Considere a água, diz ele. Nós todos estamos familiarizados com o comportamento da água no macroplano habitual dos objetos cotidianos. Nós conhecemos sua sensação de umidade, sa-

bemos que, se pura, é inodora, que pode ser bebida, que adota a forma do interior do recipiente onde é contida e assim por diante. Mas por que a água se comporta dessa maneira? A resposta deve ser encontrada, diz Searle, no microplano, quando nós entendemos que água é um composto de milhões de moléculas de água, individualmente invisíveis, feitas de dois átomos de hidrogênio e um átomo de oxigênio ( $H_2O$ ). O mesmo é verdadeiro no caso dos sólidos. Mais uma vez, o comportamento superficial de algo como o carbono pode ser entendido em termos de sua microestrutura. O diamante e o grafite são ambos compostos de átomos de carbono, mas enquanto o diamante é a substância mais dura de que temos conhecimento, o grafite é macio e escorregadio. Por quê? Porque os átomos no grafite estão arranjados na forma de uma lâmina, permitindo que uma lâmina escorregue suavemente sobre a outra, enquanto os átomos de diamante são arranjados como um octaedro, formando uma estrutura rígida que não pode ser facilmente penetrada.

Como vimos, Searle sustenta que esses aspectos superficiais, macro, são causados pelo comportamento dos microelementos subjacentes. É sobre isso que eu devo registrar meu primeiro desacordo. Se nós aprendemos alguma coisa com a análise de Hume da causação (veja Capítulo 4, seção "Behaviorismo analítico") é que uma relação causal apenas pode prevalecer entre o que ele chama de "existências distintas" entre o que podemos chamar de particulares discretos. Esses particulares não precisam ser substâncias lógicas, ainda que substâncias lógicas sejam certamente aptas, mas elas precisam incluir distintos estados de coisas. Então o sorriso em meu rosto, apesar de não ser uma substância lógica, mas uma modificação de tal substância, pode lhe causar um sentimento de alegria. Sua felicidade

também não é uma substância lógica, é um de seus estados, mas é um estado de coisas discriminavelmente diferente do ponto de vista numérico. Eu poderia sorrir sem que você se sentisse feliz, e você poderia se sentir feliz mesmo se eu não sorrisse. Os dois estados de coisas poderiam existir independentemente um do outro, e consequentemente podem ser identificados independentemente um do outro.

Tendo em mente a análise de Hume da causação, veja agora o Exercício 6.1 antes de prosseguir sua leitura.

### **EXERCÍCIO 6.1**

Você pode prever qual objeção pode ser levantada à alegação de Searle de que o comportamento da microestrutura da água causa seu comportamento no macroplano?

A objeção é que a relação entre o comportamento na microestrutura e o comportamento no macroplano é muito próxima para ser causal. Parece-me que o comportamento no macroplano é idêntico ao comportamento das moléculas no microplano. Não existem dois fenômenos aqui, o microcomportamento e o comportamento de um fenômeno numericamente diferente no macroplano, mas um único fenômeno observado sob duas perspectivas. Imagine-se observando a olho nu a água ondulada em um tanque. Você então começa a observar a mesma água através de um microscópio e ela parece completamente diferente. Suponha que a ampliação possa aumentar sem limites. Chegaria então a um ponto, em teoria, onde você não mais veria as ondações, mas as moléculas individuais nadando pelo tanque.

É verdade que há dois níveis de descrição aqui, mas alguém não pode falar de um nível de descrição causando outro nível de descrição porque descrições não

são eventos ou estados de coisas não-lingüísticos. Como parte da linguagem, as descrições podem ser relacionadas logicamente umas com as outras. Por exemplo, uma descrição pode implicar logicamente outra descrição ou logicamente contradizê-la. Se você é um filho único, por exemplo, isso deve implicar que você não tem irmãos ou irmãs. Concluir de outra forma leva a uma contradição formal. Você não pode, logicamente, ser um filho único e também possuir uma irmã (natural, não-adotada). Mas relações lógicas entre descrições, que podem ser discernidas pela pura razão independentemente da experiência, não são causais em sua natureza, porque causa e efeito são relacionadas contingentemente. Reciprocamente, relações causais podem prevalecer entre eventos ou estados de coisas, mas relações lógicas não. Causa e efeito são relacionadas contingentemente como uma questão de fato e são discerníveis apenas com base na experiência. Se uma causa e um efeito parecem ser logicamente relacionados, isto é apenas em virtude de como eles são descritos. Relações lógicas não podem se manter entre estados de coisas não-lingüísticos ou eventos em que cada um comprehenda apenas a metade da relação causal.

É verdade que um conhecimento da microestrutura das coisas pode ser usado para explicar suas macropropriedades, mas isso não é uma relação de causa e efeito no sentido que Hume comumente atribui à causa, como esboçado no Capítulo 4, seção "Behaviorismo analítico". Mais propriamente, nós entendemos o comportamento da água no macroplano quando entendemos como a água é composta, o que entra na sua composição. Não podemos dizer que os microaspectos causam os macroaspectos, porque eles são constitutivos dos macroaspectos. Os microaspectos apenas são os macroaspectos vistos

de muito, muito perto. Não faz sentido falar de um aspecto causando o outro, porque isso seria equivalente a dizer que um fenômeno é a causa de si mesmo, e isto é claramente uma posição absurda a adotar.

É importante ser muito claro sobre isso, porque me parece que Searle turva a água introduzindo aqui uma concepção não-padrão de causação. Searle reconhece que é correto dizer que a solidez pode ser definida como a estrutura entrelaçada da organização molecular, mas ele alega que isso não é na realidade uma objeção à análise que ele propõe. Ele então sustenta que nós podemos dizer, indiferentemente, que a solidez é apenas a estrutura entrelaçada do sistema de moléculas e que, assim definida, ela causa a resistência ao toque e à pressão, ou então que a solidez consiste em um nível mais elevado de aspectos tais como a rigidez e a resistência ao toque, e que isso é causado pelo comportamento dos elementos no microplano.

Eu preciso discordar imediatamente. O aspecto macroscópico da solidez, e o aspecto que consiste na estrutura entrelaçada do comportamento das moléculas no microplano, são ambos o mesmo aspecto descrito em diferentes termos. Consequentemente, o microaspecto não pode causar o macroaspecto, porque causas precisam ter existências distintas. Eu vou retornar a essa observação posteriormente, no contexto da relação mente/corpo.

A segunda expressão que Searle costuma utilizar para caracterizar a relação do mental com o físico – "realizado no" – é, como eu observei, enigmática. Searle explica que, no nível da molécula individual de água, você não vai encontrar a liquidez. Você não pode dizer de moléculas individuais que elas são líquidas. A liquidez apenas pode ser encontrada quando há um agregado de moléculas. Porém, isso não é realmente surpreendente quando a liquidez

é apenas o comportamento de um agregado de moléculas visto do macroplano.

O que dizer sobre a consciência? De novo, como Searle é perspicaz em enfatizar, você não pode apontar para um único neurônio e dizer "isto é consciente". E sobre o comportamento de toda uma coleção de neurônios, todo o cérebro em funcionamento? Nesse caso, um pensamento momentâneo revela que a analogia com a água colapsa. Quanto mais atentamente você examina o cérebro, mais você verá sobre o movimento de moléculas e transações químicas, mas você não vai subitamente deparar-se com a consciência. Isso é assim porque, como eu disse muitas vezes nos capítulos anteriores, a consciência não parece ser um fenômeno público igualmente acessível para todos os observadores, mas algo radicalmente privativo do indivíduo cuja consciência ele é, algo incluído a partir de uma perspectiva puramente subjetiva de primeira pessoa. O comportamento das moléculas, por outro lado, é algo público, que pode ser apreendido pelo ponto de vista da terceira pessoa.

É por isso que eu acho a afirmação de Searle, de que fenômenos mentais são realizados na estrutura do cérebro, enigmática. Isso é algo que ele afirma frequentemente nos primeiros escritos sobre a mente em seus livros *Intentionality* (1983) e *Minds, Brains and Science*\* (1984). Por exemplo, ele diz no primeiro desses livros que a sensação de sede está localizada no hipotálamo e que as experiências mentais estão bem lá no cérebro onde a descarga de um grande número de neurônios, em literalmente milhões de sinapses, as realizou.<sup>8</sup>

Nós poderíamos examinar as operações do hipotálamo de Searle tanto quanto queiramos, assim como os processos em seu córtex visual (a parte do cérebro

associada com a visão, ou responsável por ela), mas nós não iremos nos deparar com a sede ou com as experiências visuais de Searle. Aliás, supor que nós poderíamos encontrar experiências visuais, literalmente, na cabeça de Searle é cometer a falácia do homúnculo. Isso seria como postular um Searle miniaturizado que observa uma tela de televisão neural em sua cabeça.

Se necessário, veja de novo o Capítulo 3, seção "O mental e o espacial", para refrescar sua memória sobre a diferença entre estados mentais e processos mentais, e então veja o Exercício 6.2.

### **EXERCÍCIO 6.2**

Qual o outro erro que Searle cometeu tentando localizar a sede e as experiências visuais no cérebro?

A tentativa de localizar literalmente a sede e as experiências visuais é ilustrativa de um outro erro flagrante. Ao localizar a sede, por exemplo, de modo tão preciso no hipotálamo, Searle, eu sustento, ilegitimamente elevou a sede à condição de uma substância lógica, isto é, a uma coisidade genuína. Uma vez que esse movimento reificador tenha sido produzido, torna-se legítimo questionar onde exatamente no espaço está a coisa, ou o agrupamento de coisas. Se as coisas em questão são neurônios, a exigência em saber sua localização precisa é admissível. Isso se deve ao fato de que eles se tornam particulares básicos, substâncias lógicas, e suas identidades, ao contrário das identidades de grandes e pequenos sorrisos, não dependem logicamente de nada mais. Um sorriso grande ou pequeno, por comparação, é um particular logicamente dependente. Ele é o sorriso particular grande ou

\*N. de T.: Em português: *Intencionalidade e Mentes, cérebros e ciência*.

pequeno que é em virtude da identidade da pessoa à qual o sorriso grande ou pequeno pertence.

O mesmo é verdadeiro no caso de experiências visuais ou de sede. Como particulares logicamente dependentes, devem suas identidades às pessoas que estão tendo experiências visuais ou estão sentindo sede. Obviamente nós não podemos nos deparar com essas experiências ficando a esmo, e desligado a respeito do mundo e depois ter de olhar para ver quem as possui. Em vez disso, nós aprendemos a falar sobre sede e experiências como estados ou condições de pessoas vivas. Mas então, concebidas dessa maneira, a única resposta para a questão "Onde está a sede da pessoa?" é dizer que ela está onde quer que a pessoa esteja, em uma caminhada por um país quente, por exemplo. A sede, como um estado da pessoa, não pode ser localizada de forma mais precisa do que essa.

Se nós não compreendemos esse ponto, então nós poderíamos conjecturar que minha sede está, por exemplo, na minha garganta. Mas embora seja verdadeiro dizer que eu sinto sede porque minha garganta se sente ressecada ou seca, seria muito estranho dizer que eu sinto sede em minha garganta, e ainda mais estranho seria dizer que minha garganta sente sede. Sou eu que sinto sede porque eu não bebi água, e como resultado minha garganta se sente seca. Para não cometer erros, sou eu, a pessoa, que tem um sentimento de segurança na sua garganta. A garganta por si mesma não sente nada. Como ela poderia, se não é uma pessoa, mas apenas uma parte de minha anatomia? O mais estranho de tudo seria localizar a sede literalmente em meu cérebro, como Searle faz. Isso iria resultar na absurdade de que meu cérebro sente sede.

Dizendo tudo isso eu posso ser acusado de ser injusto com Searle. Quando ele

diz que a sede e as experiências visuais se dão na microestrutura do cérebro, isso poderia ser meramente uma maneira de afirmar que esses fenômenos mentais são em realidade apenas fenômenos físicos, que afinal de contas esses estados consistem em nada mais que o comportamento de neurônios descritíveis em um vocabulário puramente fisicalista. Porém, nós podemos dizer a partir da leitura de Searle que isso não pode ser sua real intenção, pois ele explicitamente rejeita qualquer comprometimento com o materialismo redutivo, como nós veremos em breve.

Esse repúdio ao materialismo redutivo, especificamente nos primeiros trabalhos de Searle, não foi, talvez, sempre tão claro quanto poderia ter sido. As afirmações de que sede e experiências visuais estão bem lá no cérebro poderiam, como nós testemunhamos, terem sido interpretadas como uma defesa indireta da identidade do mental com o físico. Porém, quando *Searle and His Critics*\* foi publicado, em 1991, o próprio Searle estava deixando muito claro seu repúdio ao materialismo radical.<sup>9</sup> Nesse livro, em uma resposta ao teórico da identidade australiano David Armstrong, Searle escreveu (e é essencial que eu aqui faça uma citação na íntegra para deixar o ponto bem esclarecido):

Armstrong pensa que, se uma propriedade é causalmente superveniente do comportamento dos elementos de sua microestrutura, então isto também deve ser ontologicamente redutível. Em sua visão, propriedades supervenientes de nível mais alto como a solidez e a consciência, porque elas estão sujeitas ao tipo de redução causal envolvida na superveniência causal, são também, por essa razão, ontologicamente redutíveis. Para Armstrong solidez, consciência e Intencionalidade estão todas no mesmo

\*N. de T.: Em português: *Searle e seus críticos*.

barco que o calor: não há nada lá além do comportamento dos microelementos. Agora, enfaticamente, essa não é a minha visão, e eu acredito que isso está equivocado. A superveniência, ainda que estrita, não implica reducionismo. Do fato de que uma propriedade é superveniente ao comportamento dos elementos do nível inferior simplesmente não se segue que não há nada lá exceto o comportamento dos elementos do nível inferior. No caso, por exemplo, da consciência, nós temos uma propriedade superveniente, mas, não obstante, não-redutível. Isso, de fato, é a diferença entre, por exemplo, o calor e a consciência. A consciência é uma propriedade separada e não-redutível, e em um sentido é, eu suponho, emergente. De fato, concordando comigo que certas características tais como a liquidez e a solidez são emergentes, me parece que Armstrong está comprometido em negar sua redutibilidade ontológica, porque dizendo que a propriedade é emergente nós estamos dizendo que há um fenômeno separado aqui, ainda que o fenômeno possa ser inteira e causalmente superveniente de outros fenômenos. Redução causal, sim; redução ontológica, não.<sup>10</sup>

Essa passagem deixa amplamente claro o comprometimento de Searle com o materialismo não-redutivo. Propriedades mentais como a consciência e a intencionalidade são ontologicamente distintas de propriedades físicas como o calor, e não redutíveis ontologicamente ao comportamento de microelementos puramente físicos, como em uma teoria cinética dos gases, por exemplo. Na opinião de muitos filósofos, isso compromete Searle com a aceitação de um dualismo de propriedades, à visão de que, assim como suas propriedades físicas, como cor, massa e extensão, o cérebro possui o aspecto não-físico da consciência.

Porém, Searle nega vigorosa e repetidamente que ele seja um dualista de propriedades. Aceitar o dualismo, até mesmo o dualismo de propriedades, é, de acordo com Searle,

negar a visão científica sobre o mundo que nós dolorosamente alcançamos ao longo de muitos séculos passados. Aceitar o dualismo é concluir que nós realmente vivemos em dois mundos muito diferentes, um mundo mental e um físico, ou ao menos que existem duas espécies de propriedades, mentais e físicas. Eu espero ter deixado claro [...] que penso que alguém pode aceitar a existência e a irredutibilidade da consciência como um fenômeno biológico sem aceitar a ontologia do dualismo tradicional, a ideia de que há duas espécies diferentes, metafísica ou ontologicamente, de domínio em que vivemos, ou duas espécies diferentes de propriedades no mundo.<sup>11</sup>

Essa passagem, que é tomada quase do final do *The Mystery of Consciousness\**, muito dificilmente se ajusta, porém, com a bem disposta aceitação de Searle no início do livro:

Nós não temos nada que se assemelhe a uma ideia clara sobre como os processos cerebrais, que são publicamente observáveis, fenômenos objetivos, poderiam causar qualquer coisa tão peculiar quanto estados qualitativos internos de consciência ou sensibilidade, estados que são em algum sentido “privados” ao possuidor do estado [...] Como poderiam esses fenômenos qualitativos privados, qualitativos, serem causados por processos físicos correntes como descargas neuronais eletroquímicas nas sinapses de neurônios?<sup>12</sup>

O tom dessa passagem, talvez contrariando as intenções de Searle, é para dar

\*N. de T.: Em português: *O mistério da consciência*.

a entender que há uma falha explicativa, incapaz de ser coberta ou fechada, entre o que ocorre física e publicamente e o que ocorre mental e privadamente. Se o mental fosse físico, então presumivelmente não seria radicalmente privado e subjetivo. Realmente, uma parte integrante do que significa classificar alguns fenômenos como físicos, em oposição a não-físicos, é certamente que ele é público, objetivo e não perceptível a partir de uma perspectiva subjetiva puramente de primeira pessoa. Se a posição de Searle não é o dualismo de propriedades expresso em outros termos, então eu temo não saber o que ela é. Não obstante o fato de que ele continuadamente repita que a consciência e a intencionalidade são fenômenos biológicos, naturais – uma posição muito razoável à luz das dificuldades com um dualismo de substâncias plenamente desenvolvido –, eu acredito que Searle deve graciosamente aceitar que ele é, afinal, um dualista de propriedades.

Uma virtude da versão de Searle do monismo não-reduutivo é que ele deixa bem claro que a razão pela qual a relação de superveniência se sustenta é porque o físico causa o mental. De fato, Searle sugere que o termo “superveniência”, e a ideia de qualquer *nova* relação que isso denota, pode ser prescindido, porque não mais cumpre função alguma na filosofia uma vez que a existência da causação de baixo para cima, de micro para macro, é reconhecida:

Os aspectos formais da relação já estão presentes na suficiência formal das formas de causação micro/macro. E a analogia com a ética é apenas uma fonte de confusão. A relação dos macroaspectos mentais do cérebro com seus aspectos microneuronais é totalmente diferente da relação da bondade com os aspectos do fazer o bem, e causa confusão tratá-los conjuntamente.<sup>13</sup>

Se Searle está certo, e a postulada superveniência dos aspectos mentais sobre os físicos é na realidade apenas outra maneira de falar sobre a dependência causal da mente sobre o corpo, então o exame dos pontos fracos e fortes dessa teoria se torna mais claro. A afirmação de Searle de que os processos físicos causam ocorrências mentais parece ser eminentemente razoável, e é com base nisso que continuarei a fazer minha avaliação do dualismo de propriedades como uma teoria da mente.

### Três problemas para o dualismo de propriedades

O dualismo de propriedades tem três grandes problemas. Eles se apresentam da seguinte forma:

1. Qual é a natureza das propriedades mentais não-físicas que coletivamente compõem nossas vidas mentais e que supostamente são suscitadas por processos físicos no cérebro?
2. Em que medida nós podemos explicar, e entender, como processos físicos no cérebro dão origem a estados não-físicos da mente?
3. Como pode o dualismo de propriedades, e a dependência causal da existência do mental sobre o físico, evitar uma ida direta ao epifenomenalismo, à afirmação de que a causação segue em uma via apenas, do cérebro para a mente, e não da mente para o cérebro, então fazendo da mente uma espectadora impotente e que não contribui ao que ocorre no mundo?

No restante deste capítulo tentarei lidar com as primeiras duas questões. A terceira, que nos remete ao problema de Descartes sobre como a mente, concebida

como uma alma, pode afetar o corpo, requer um tratamento mais extenso, como veremos.

*Qual é a natureza das propriedades mentais?*

Nossa primeira questão indagou: qual é a natureza das propriedades mentais que coletivamente compõem nossas vidas mentais?

Aqui, nós somos levados diretamente de volta a Descartes, que certamente estava correto ao sustentar que um aspecto essencial de nossa vida mental é a consciência, nossa consciência das coisas à nossa volta, assim como de nossos estados internos. Completamente isentadas de qualquer vestígio da consciência, nossas vidas mentais efetivamente deixaram de existir. Ainda que, como salientamos, possam existir crenças, desejos, intenções e assim por diante das quais circunstancialmente não nos apercebamos, esses estados podem, quando necessários, ser trazidos de volta à consciência, e esta pronta acessibilidade à consciência constitui um aspecto integral deles.

Mas responder que a consciência é essencial para a própria existência de nossos estados mentais ainda não nos diz o que a consciência é. Então nossa próxima questão tem de ser: qual é a natureza da consciência ou do próprio estar consciente? Respostas a essa questão incluíram observações tais como "Consciência é como jazz; se você tem de perguntar o que é, você nunca saberá", ou "Consciência é como a Trindade; se você pode explicar o que ela é, você não a compreendeu". Eu penso que o ponto sério por trás dessas observações é que em um sentido todos nós sabemos o que a consciência é, simplesmente pelo fato de que nós somos conscientes. Nós todos sabemos o que é ver, escutar, cheirar, tocar e sa-

borear as coisas; experimentar sensações e emoções; ter pensamentos e recordar. Em síntese, nós sabemos que há algo que é como nós [somos], para usar a frase de Thomas Nagel,<sup>14</sup> e neste sentido nós sabemos o que é desfrutar ou aguentar a consciência. Se, por outro lado, nós não sabemos o que é ser consciente, não há alguém que nos possa dizer, e, ademais, seria sem sentido para qualquer um tentar, pois uma condição essencial para seu sucesso é que nós já estejamos conscientes; de outra maneira nós não poderíamos ter qualquer consciência ou entendimento do que eles estavam tentando nos dizer. O que ainda há para ser dito sobre o que a consciência é, e de que maneira, na condição de seres que expericiam, poderíamos deixar de conhecer o que ela é?

Mas, será sem dúvida se objetará, a questão pediu não por *exemplos* da consciência, mas por uma *definição* dela. Em outras palavras, a demanda foi de que deveríamos ser capazes de explicar detalhadamente em outros termos que não contenham, explícitos ou implícitos, o termo "consciência" ou seus equivalentes, precisamente o que a consciência é. Isso pressupõe que uma análise redutiva sobre no que consiste a consciência é possível. O perigo é, porém, que se nós aceitarmos essa pressuposição e depois falharmos em apresentar detalhes convincentes do que comprehende a consciência, expressos em outros conceitos, pareceria então que nós teríamos falhado em entender propriamente o que a consciência é.

Agora veja o Exercício 6.3 antes de prosseguir a sua leitura.

#### **EXERCÍCIO 6.3**

É possível definir consciência? Você pode fornecer uma definição não-circular da consciência que satisfaça à fórmula:

**mula: Um ser pode ser verdadeiramente descrito como consciente se, e somente se,?**

Minha resposta é rejeitar a pressuposição de que uma redução analítica da consciência é possível. (Dada a força principal do argumento deste livro, não é preciso dizer que eu não penso que uma redução ontológica seja possível tampouco.) Eu não acredito, apesar de não ver como poderei provar rigorosamente a alegação para alguém que permanece hesitante, que a consciência, para adaptar as palavras do filósofo Bishop Butler (1692-1752), é o que é e não outra coisa qualquer. A noção de consciência é logicamente primitiva ou básica, isto é, ela não pode ser subdividida em termos ou componentes mais simples, mais básicos. Consequentemente, não é possível fornecer uma análise não-circular do que a consciência é, e isso é equivalente a dizer que nenhuma análise é possível, sob qualquer condição. Termos sinônimos podem ser fornecidos, e muitos exemplos de estados nos quais a consciência é parte essencial para a sua própria existência podem ser dados, como nos casos apresentados, mas não parece ser possível decompor a consciência em estados de coisas mais básicos, ou fornecer uma descrição mais básica do que a consciência é, usando outros termos.

*Como o cérebro físico suscita aspectos mentais?*

A questão do que a consciência é – para a qual sugeri que nenhuma resposta do tipo petição de princípios será ou poderá vir a ser apresentada – deve ser mantida completamente distinta de como a consciência é causada. Mesmo que nós não possamos explicar detalhadamente em outros termos o que a consciência é,

isso não é razão para concluir que nós não seremos capazes de um dia explicar como a consciência foi produzida pelos mecanismos físicos que a sustentam. De fato, devido a todos os detalhes que nós descobrimos sobre como o cérebro funciona, é inteiramente razoável pensar que eventualmente seremos capazes de fornecer as causas físicas subjacentes da consciência, e então entender como ela depende do físico para sua existência.

Porém, há filósofos que irão imediatamente me acusar de um otimismo equivocado ao fazer essa afirmação. De acordo com essa maneira de pensar, mesmo se eu conseguisse identificar esses eventos cerebrais cujas ocorrências parecem ser essenciais para a existência da consciência, eu ainda não teria realmente explicado como a consciência é suscitada pelos processos físicos no cérebro. Dizendo isso de outra maneira, suponha que eu possa ligar e desligar os processos físicos no cérebro dos quais a consciência parece depender. Eu desligo o processo, e a pessoa cai inconsciente. Eu os ligo de novo, e a pessoa se recupera e desperta. Em todo o mundo, vamos supor, vários cientistas põem essa teoria sob teste, e todos concordam que alguns eventos mentais complexos, descritos em detalhes físicos e químicos muito precisos, pareceriam ser essenciais para a existência da consciência. Bem, de acordo com a objeção que eu estou em vias de esboçar, mesmo se tudo isso ocorresse, nós ainda não teríamos alcançado um entendimento apropriado e uma explicação de como o físico produz a consciência.

A base da objeção é esta: ocorrências físicas não apenas parecem ser diferentes da consciência; elas são absolutamente diferentes, tão absolutamente, de fato, que é inconcebível como o físico poderia produzir o mental. Não é difícil simpatizar com essa posição. Eu penso que seria honesto dizer que nós possuímos uma

intuição forte, predominante, de que há um abismo permanente entre o material e o mental, que não apenas proíbe para sempre suas identificações, mas, além disso, coloca fora de questão uma explicação de como um dá origem ao outro. Como a água do cérebro transforma-se no esplêndido vinho da consciência, para tomar emprestada a admirável metáfora de Colin McGinn.<sup>15</sup> A acumulação de detalhes neurofisiológicos, ao que parece, nunca tornará transparente sua relação com a mentalidade, muito menos demonstrará que a consciência e a neurologia são dois nomes para um único fenômeno.

McGinn, de forma pessimista, acredita que nós nunca decifraremos o problema de como a consciência emerge da atividade cerebral. Ele argumenta que nossa constituição sensória e cognitiva sempre nos impedirá de obter a compreensão necessária sobre como o nó mente/corpo é amarrado. Nossa falha em entender como a consciência está relacionada com eventos neurais é uma falha em princípio por causa dos limites de nossa experiência e das limitações correspondentes a respeito dos conceitos que podemos adquirir. Como resultado, é provável que sejamos para sempre impedidos de resolver o problema mente/corpo.

A teoria de Hume sobre como formamos conceitos nos ajuda a ilustrar a abordagem de McGinn. De acordo com Hume, conceitos são supostamente imagens pálidas derivadas de impressões sensoriais mais vivas. Nossos sentidos respondem apenas a uma pequena parte do espectro eletromagnético – nós não podemos ver ondas de rádio, por exemplo. Então nós não possuímos, na teoria de Hume, um conceito genuíno da cor das ondas do rádio, porque não há impressão do sentido a partir da qual derivar o conceito apropriado. Por outro lado, nós podemos ver rosas, então nós certamente sabemos o que

se quer dizer quando descrevemos uma rosa como vermelha. Argumentando analogicamente, podem existir outros aspectos do universo que nós não estamos em posição de detectar, mas, *ex hypothesi*, nós não temos a possibilidade de dizer o que eles são. Não sendo capaz de ter experiência desses aspectos, nós não poderíamos, em uma teoria empirista simples como a de Hume, formar qualquer conceito deles. Mas, então, qualquer teoria explanatória que empregue esses conceitos não seria acessível a uma mente humeana. Conduzindo essa linha de argumento para sua conclusão, pode-se dizer que podem existir criaturas capazes de formar conceitos e teorias não disponíveis a nós, porque seus aparatos de percepção são diferentes dos nossos e/ou porque eles são em geral cognitivamente diferentes de nós.

A menos que acreditemos em mágica e no miraculoso, presume-se que deve haver *alguma* teoria que possa explicar como estados conscientes são causados por estados mentais. Mas poderia ser o caso – e de fato a própria intratabilidade do problema mente/corpo e a dificuldade de conceber que aspecto uma explicação da consciência, em termos de processos mentais, poderia concebivelmente ter, sugere que isto de fato é o caso – de que essas teorias são simplesmente indisponíveis para nós porque elas empregam conceitos que não podemos adquirir. Então, segundo McGinn, é possível que haja alguma propriedade no cérebro que nos capacitaria a explicar naturalisticamente a consciência, com a condição de que pudéssemos compreender a natureza desta propriedade. Porém, se tal propriedade existisse, McGinn sustenta, parece que haveria apenas duas avenidas que poderiam nos dar acesso a ela.

Uma é por meio de nossa própria experiência da consciência. Infelizmente, como McGinn salienta, a consciência

não parece ser capaz de ajudar, porque a reflexão sobre nossos próprios estados fenômenicos, ou o exame deles, não fornece indício de qualquer espécie sobre como tais estados podem depender do cérebro, nem dele surgir. Nós todos sabemos o que é ser consciente de maneira bastante independente de qualquer conhecimento do que ocorre em nossos cérebros, ou mesmo sem saber que temos cérebros. A reflexão sobre o fato da nossa experiência consciente não nos abre, por assim dizer, janela alguma para qualquer processo físico que pudesse estar envolvido em dar origem a ela, ou na sua constituição.

A outra via possível para a compreensão ocorre por meio da pesquisa do cérebro e da neurociência em geral. Mas aqui de novo a loteria nos sai em branco, essencialmente pela mesma razão dada por Leibniz (veja a "Introdução" ao Capítulo 1). Eu posso olhar em seu cérebro, mas eu não serei, por meio disso, capaz de discernir o que você está pensando: tudo que verei, nas palavras de Leibniz, são "partes que empurram e movem uma à outra, e nunca alguma coisa que possa explicar a percepção".<sup>16</sup> Acrescentando à observação de Leibniz, McGinn escreve:

Porque os sentidos são ajustados para representar um mundo espacial [...] eles essencialmente apresentam as coisas no espaço com propriedades espaciais definidas. Mas são precisamente *tais* propriedades que parecem inerentemente incapazes de resolver o problema mente/corpo: nós não podemos vincular a consciência ao cérebro em virtude das propriedades espaciais do cérebro [...] A consciência não parece ser feita a partir de processos espaciais menores; no entanto, a percepção do cérebro parece limitada a revelar tais processos.<sup>17</sup>

David Chalmers concorda com McGinn:

[O problema] reside em toda a estratégia de explicação. Esses modelos e teorias [físicalistas] simplesmente não são o tipo de coisa que pode explicar a consciência [...] Qualquer explicação dada em termos puramente físicos [...] irá apenas produzir [...] uma estrutura e uma dinâmica. Embora isso seja suficiente para dar conta da maioria dos fenômenos naturais, o problema da consciência vai além de qualquer problema sobre a explicação da estrutura e função; assim sendo, uma nova espécie de explicação se faz necessária.<sup>18</sup>

Não há como negar que as sugestões de McGinn e Chalmers podem exercer uma grande atração. Dada a total disparidade dos fenômenos em questão – os processos cerebrais públicos, objetivos, extensos que exemplificam, por um lado, a matéria em movimento, e, por outro lado, modos de consciência privados, subjetivos, não-extensos, para os quais os conceitos de movimento e repouso não possuem aplicação –, certamente parece difícil entender como o essencialmente mecânico pode dar origem ao essencialmente não-mecânico.

Contudo, parece-me que McGinn e Chalmers estão errados ao visualizar a questão dessa maneira. A relação mente/corpo é difícil de compreender, de fato, mas a abordagem adotada por esses dois pensadores torna a sua compreensão ainda mais difícil do que o necessário. Em síntese, o modo como eles conceitualizam a dificuldade coloca a compreensão do problema permanentemente fora do alcance, transformando um problema difícil em um mistério insolúvel.<sup>19</sup>

O que motiva a abordagem de McGinn é o desejo, ao que me parece, de tornar cristalino como o físico dá origem ao mental. Dizendo dessa maneira, a pretensão parece eminentemente razoável, até que voltarmos nossa atenção ao que tal abordagem tem a pretensão de ser. Parece que McGinn está à procura de que deva-

mos, de alguma maneira, tornar transparente como processos materiais causam a mente, e isso pode ser interpretado, por sua vez, como a reivindicação de que seja mostrado como, dada a ocorrência de certas espécies de evento físico, certas espécies de evento mental *devem* ocorrer. O que ele quer, como ele mesmo confessa, é encontrar algum aspecto natural ou propriedade do cérebro que, se nos fosse possível compreender sua natureza, iria revelar-se como o mecanismo por meio do qual o movimento mecânico se transforma em pensamento consciente. No momento, a emergência da mente a partir da matéria nos parece ser um milagre, mas ao mesmo tempo sabemos que não pode ser um milagre e que deve haver uma explicação racional. Nossa dificuldade é que simplesmente não podemos entender como o truque de mágica está ocorrendo, porque nós não possuímos um *insight* da natureza do vínculo oculto entre os processos físicos e a experiência consciente.

A crença de que deve haver tal vínculo é a fonte da visão equivocada de McGinn sobre como nós vamos conseguir uma compreensão das causas físicas da consciência. O que McGinn espera alcançar, de fato, é exibir a conexão necessária entre o físico e o mental, o nexo perdido, escondido e secreto, que, bastaria apenas que fosse desvendado aos nossos olhos para revelar em um instante como o mental deve necessariamente e invariavelmente surgir da atividade física subjacente. Mas essa busca por tal nexo é uma caça a ganso selvagem. A busca por uma conexão necessária, um vínculo misterioso que mostre como o fenômeno da consciência está ligado com a atividade cerebral, e dela emerge, é infrutífera, porque tudo que um exame cada vez mais próximo de uma relação causal pode revelar são mais elos na cadeia, isto é, mais eventos que são causalmente relacionados uns com os

outros. Um engano categorial está em evidência aqui. A relação causal não é em si mesma um evento ou um objeto. Porém, muitos eventos são descobertos em uma cadeia causal, e nenhum, em princípio, pode ser considerado como aquele que estamos procurando. No final, tudo que descobriremos é que certas espécies de eventos seguem com regularidade outras espécies de eventos, e que uma verdadeira afirmação descrevendo uma relação causal entre um evento do tipo A e um do tipo B vai implicar um condicional contrafactual, que afirma que se A não ocorrer, então B tampouco ocorrerá. Qualquer conexão necessária e misteriosa entre os eventos, o vínculo que os une, não apenas não poderia ser encontrada, mas em princípio nunca poderá sê-lo. A busca seria recursiva. Seria como procurar por um substrato sem traços característicos, que subordinasse todas as qualidades perceptíveis. O que quer que seja com que a busca tenha se deparado iria se mostrar ser apenas outra qualidade perceptível, anteriormente não observada. Exatamente da mesma maneira, o nexo oculto que estamos buscando, que supostamente esclarece como o físico causa o mental, e torna completamente transparente e inteligível como um é o resultado do outro, iria apenas recuar a nossa frente como um engodo.

O que tanto McGinn como Chalmers veem como inaceitável é que possa vir a ser o caso de que processos físicos complexos apenas dêem origem a estados conscientes, e que não haja mecanismo oculto algum se mostrando como propriedades do cérebro que estão por ser descobertas, mas ainda não foram, que tornariam transparentes o vínculo causal entre a matéria e a mente. Essa inabilidade em aceitar que não há mecanismo oculto algum se manifesta, no caso de Chalmers, em sua afirmação na passagem citada an-

teriormente, de que “Qualquer explicação dada em termos puramente físicos [...] irá apenas produzir [...] uma estrutura e uma dinâmica”. Mas por que isso deve ser assim? O que Chalmers está afirmado é que um efeito deve ser como sua causa para que a relação causal possa ser possível e qualquer explicação que a utilize seja inteligível. A suposição não declarada é a de que, se a consciência tem um caráter totalmente diferente dos processos cerebrais, então ela não pode resultar de tais processos, e que deve haver algum vínculo mediador misterioso entre o físico e o mental que explica como o último pode resultar do primeiro. Mas não há garantia para essa conclusão, como uma famosa passagem de Hume demonstra:

Se você finge [...] provar que [...] uma posição de corpos nunca pode causar pensamento; porque, virando como você quiser, isto não é nada além de uma posição de corpos; você deve, pela mesma linha de raciocínio, concluir que isto nunca pode produzir movimento; pois não há conexão mais aparente em um caso que no outro.<sup>20</sup>

Em outras palavras, antes da experiência, qualquer coisa pode ser a causa de qualquer coisa. Uma pessoa poderia, por um simples desejo, aniquilar o Sol ou controlar os planetas em suas órbitas. O mundo não funciona dessa maneira, mas não há contradição em supor que ele poderia assim funcionar. Similarmente, processos físicos complexos poderiam dar origem ao pensamento e à consciência; a possibilidade não pode ser descartada antecipada e independentemente da experiência de como o mundo funciona efetivamente. De fato, de tudo que nós sabemos sobre a maneira como a sofisticação mental se combina com a complexidade neurológica, na medida em que ascendemos na escala evolucionária, junto com uma ri-

queza de evidências bem documentadas dos efeitos sobre a mentalidade causados pelos danos ao cérebro, ou o uso de drogas, é mais do que razoável supor que o físico dê origem ao mental. Como ele consegue isso, porém, não é mais misterioso, em última análise, do que o modo pelo qual *qualquer* causa dá origem a seu efeito. Se a causalidade do físico para o mental parece inexplicável, então igualmente deve ser assim com a causalidade do físico para o físico. Não há, como Hume salientou, mais conexão discernível em um caso que no outro. Visto que, admito, McGinn e Chalmers não estão especialmente interessados em tentar penetrar o segredo de como um evento físico leva a outro, eles não têm motivo para procurar por algum ingrediente misterioso que iria supostamente revelar como o físico traz o mental à existência. Não há, como eu tenho tentado explicar, tal ingrediente a ser encontrado, e exigir que tal ingrediente deva ser produzido é um ideal falso do que significa explicar um fenômeno em termos de outro. É o desejo, o anseio por um *insight* especial sobre como o cérebro produz a mente, o erro que motiva as abordagens de McGinn e Chalmers, pois não há tal *insight* especial a ser obtido, a não ser descobrindo, por meio de investigação científica cuidadosa e paciente, de quais estados físicos do cérebro a consciência realmente depende.

Se compararmos a causalidade físico/mental com a causalidade físico/físico, não poderemos reclamar por termos alguma compreensão especial de como um evento físico dá origem a outro evento físico, o que carecemos no caso de como um evento físico dá origem a um evento mental. Isso se deve ao fato de que tudo o que logramos observar, mesmo no caso da causalidade físico/físico, é certo tipo de evento seguindo outro tipo de evento e exemplificando uma dependência con-

trafactual sobre ele. Nós comumente experienciamos um evento físico suscitando outro evento físico, como quando uma pedra quebra uma janela. Mas, igualmente, todos sabemos, por nossa própria experiência, que a causação do físico para o mental ocorre, como quando um alfinete nos faz sentir dor, ou avistar um touro em um campo que estamos cruzando causanos um aumento súbito de adrenalina e nos faz correr para a porteira.

Explicações de como uma coisa causa outra coisa devem eventualmente chegar a um fim. Nós não podemos continuar citando vínculos em uma cadeia causal. Mais cedo ou mais tarde vamos atingir o último elo, e se é então questionado por que o evento em questão foi produzido por outro evento, seremos forçados a admitir que seja simplesmente assim que as coisas acontecem. Isso se aplica igualmente para ambas, à causação de físico para físico e à causação de físico para mental. Uma compreensão de como certos processos cerebrais produzem certos estados mentais pode, em última instância, resultar em nada mais do que isso. A busca por um nexo entre a mente e a matéria é uma busca por uma quimera, e leva tão somente à mistificação.

### Conclusão

Na última parte deste capítulo eu li-dei com a questão concernente a em que medida podemos explicar como a consciência, e os estados mentais em geral, são causados pelos processos físicos. Sugeri, utilizando lições da abordagem de Hume da causação, que não há razão para pensar que nossa compreensão dessa questão precisa, em última análise, ser de algum modo inferior à compreensão que nós temos no caso das explicações causais de uma variedade de ocorrências em que a mentalidade não é envolvida.

Mas a teoria de que o cérebro causa a mente ainda não está ainda sã e salva. Alguns filósofos têm pensado que, se a mente é produzida pelo cérebro e depende dele para sua existência, então isso deve inevitavelmente levar a um epifenomenalismo inaceitável, à visão de que a mente é impotente e constitucionalmente incapaz de produzir uma diferença causal ao mundo. Uma discussão e avaliação crítica desse problema constituem o conteúdo do próximo capítulo.

### Questões para reflexão

1. Qual é o significado de "superveniência"? Ela conduz diretamente ao epifenomenalismo?
2. A consciência pode ser definida?
3. Que grau de plausibilidade você atribui à alegação de Searle de que estados mentais são causados pelo cérebro e nele realizados?
4. Um monismo não-redutivo deve negar que a existência desencarnada é logicamente impossível?

### Sugestões para leituras aprofundadas

- KIM, Jaegwon. *Mind in a Physical World*. Cambridge, MA, MIT Press, 1998 examina em detalhes o materialismo redutivo e o problema da causação mental.
- KIM, Jaegwon. *Supervenience and the Mind*. Cambridge, Cambridge University Press, 1993 vai dizer-lhe tudo que você sempre quis saber sobre variedades de superveniência. O livro inclui discussões muito profundas e rigorosas de um dos principais especialistas em superveniência.

- TYE, Michael. *Ten Problems of Consciousness*. Cambridge, MA, MIT Press, 1995. Para discussões animadas sobre as principais fontes da consciência.
- CHALMERS, David. *The Conscious Mind*. Oxford, Oxford University Press, 1996 é uma leitura essencial, incluindo discussões muito detalhadas de todos os principais temas, mas é um tanto longo e complexo.
- FLANAGAN, Owen. *Consciousness Reconsidered*. Cambridge, MA, MIT Press, 1992 é um livro intenso e acessível que cobre muitos dos temas importantes que são atualmente debatidos.
- DENNETT, Daniel. *Kinds of Minds*. London, Weidenfield and Nicolson, 1996 traz uma abordagem da emergência da mente a partir de uma perspectiva evolucionária.

## NOTAS

1. DAVIDSON, D. *Mental Events*. 1970, reimpresso em: DAVIDSON, D. *Essays on Action and Events*. Oxford, Oxford University Press, 1980, p. 214.
2. KIM, J. *Mind in a Physical World*. Cambridge, MA, MIT Press, 1998, p. 5-6.
3. Se a forma do desaparecimento da teoria da identidade tipo-tipo da mente/cérebro fosse verdadeira, soaria estranho dizer que o mental sobrevém ao físico, assim como seria equivalente a afirmar que um único fenômeno, caracterizável tanto como mental quanto físico, sobrevém a si mesmo. Talvez um teórico da superveniência pudesse tentar contornar isso, todavia, sustentando que pode não haver mudança na descrição do estado mental de uma pessoa a menos que haja mudança na descrição do seu estado físico (estado cerebral). A identidade do mental e do físico poderia ser vista como o caso que limita a superveniência. A identidade cuida dos outros aspectos da superveniência delineados anteriormente. O mental variará de acordo com o físico porque ele é físico, e dependerá do físico para sua existência, porque, de novo, ele é simplesmente o físico. O único aspecto com o qual a teoria reducionista da identidade não pode lidar é o de que a superveniência, supõe-se, seja assimétrica – o mental não pode mudar a menos que o físico mude –, mas o físico pode mudar sem que o mental mude porque estados físicos têm a possibilidade de realização múltipla de estados mentais. Se a teoria da identidade tipo-tipo fosse verdadeira, então seria igualmente o caso de que o físico não poderia mudar a menos que o mental mudasse, porque o mental é o físico.
4. HARE, R. M. *The Language of Morals*. Oxford, Oxford University Press, 1952.
5. Ibid., p. 145.
6. BURWOOD, S., GILBERT, P., LENNON, K. *Philosophy of Mind*. London, UCL Press, 1998, p. 36.
7. SEARLE, J. *The Rediscovery of the Mind*. Cambridge, MA, MIT Press, 1992, p. 1.
8. SEARLE, J. *Intentionality*. Cambridge, Cambridge University Press, 1983, p. 267.
9. LEPORE, E.; GULICK, R van. *Searle and His Critics*. Oxford, Blackwell, 1991.
10. Ibid., p. 182.
11. SEARLE, J. *The Mistery of Consciousness*. London, Granta Books, 1997, p. 194-5.
12. Ibid., p. 8.
13. SEARLE, J. *The Rediscovery of the Mind*. p. 126.
14. NAGEL, T. "What is it Like to be a Bat?" in: NAGEL, T. *Mortal Questions*. Cambridge, Cambridge University Press, 1979, p. 165-80.
15. McGINN, C. *The Problem of Consciousness*. Oxford, Oxford University Press, 1991, p. 1.
16. LEIBNIZ, G. W. "Monadology", primeiro publicado postumamente em ERDMANN, J. E., ed. *Leibniz, Opera Philosophica*. 2 v. Berlin, 1840.

17. McGINN, C. "Can We Solve the Mind/Body Problem", in: *The Problem of Consciousness*, p. 11-2.
18. CHALMERS, D. *The Conscious Mind*. Oxford, Oxford University Press, 1996, 121.
19. A acusação de mercador de mistérios atribuída a McGinn foi intitulada "O novo Misterianismo" por Owen Flanagan em seu livro *Consciousness Reconsidered*. Cambridge, MA, MIT Press, 1992, p. 9. Os Antigos Misterianos foram dualistas de substância que pensaram que a consciência seria um aspecto da alma não-física, e portanto, em princípio, além do alcance da explicação científica. Os Novos Misterianos rejeitam o dualismo de substância, mas acreditam que a natureza da ligação entre processos cerebrais e a consciência nunca será descoberta, e a relação permanecerá para sempre um mistério. Veja também McGINN, C. *Minds and Bodies: Philosophers and their Ideas*. Oxford, Oxford University Press, 1977, p. 106-7 para a revelação intrigante de que os Novos Misterianos tomaram empestado seu título de uma extinta banda de rock dos anos 60 chamada "? and the Mysterians".
20. HUME, D. *A Treatise of Human Nature*. [SELBY-BIGGE, L. A., ed. NIDITCH, P. H.; rev.] Oxford, Clarendon Press, 1978, p. 247.

# 7

## CAUSAÇÃO PSICOFÍSICA

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Estar apto a fornecer alguns exemplos de relações causais entre eventos físicos e mentais.
- Entender o que o epifenomenalismo alega e como a doutrina pode ser criticada.
- Entender por que o status de causas foi negado às razões para ações, e apreciar criticamente a resposta de Donald Davidson a essa alegação.
- Entender os principais elementos da teoria causal da ação intencional de Davidson.
- Apreciar criticamente o monismo anômalo de Davidson e entender por que se pensou que ele levaria ao epifenomenalismo.

### Introdução

Este capítulo lida com o problema da causação psicofísica. Como, em outras palavras, é possível que a mente afete o corpo? Na vida cotidiana não duvidamos sequer por um momento de que nossos estados mentais realmente afetem nossos estados físicos; contudo, explicar como isso é possível tem-se mostrado extraordinariamente difícil. A seguir gostaria de colocar o problema no contexto das recentes filosofias da mente e da ação, analisando principalmente o trabalho de

Donald Davidson nessa área e, em menor grau, o trabalho de Jaegwon Kim.

Primeiramente, fornecerei um esboço da teoria de Davidson segundo a qual desejos e crenças que são citados como razões para ações podem ser causas daquelas ações. Explicarei, depois disso, como a partir dessas suas primeiras teorizações a respeito do assunto Davidson foi levado a desenvolver uma teoria chamada de monismo anômalo – o significado dessa expressão será mais bem explicado em breve – e por que vários filósofos acreditam que essa teoria compromete Davidson com o

epifenomenalismo.<sup>1</sup> Epifenomenalismo é a tese segundo a qual estados mentais são causalmente inoperantes e, contrariando nossa visão de senso comum, a mente nunca produz uma diferença causal no mundo. Depois de tentar determinar se Davidson pode ser livrado da acusação de ser um epifenomenalista, oferecerei uma exposição e avaliação um pouco mais breve da contribuição de Jaegwon Kim à questão da causação psicofísica.

### Causação psicofísica

Comecemos recordando algumas das maneiras nas quais cremos que o mental e o físico exercem influência causal um sobre o outro. Considere as seguintes frases, que exemplificam as diferentes maneiras em que estados mentais e físicos podem ser causalmente relacionados:

1. A dor torácica intensa e repentina fez com que ele gritasse involuntariamente (*do mental ao físico*).
2. A dor em seu tórax tornou-se tão intensa que por um momento ele pensou que estava tendo um ataque cardíaco (*do mental ao mental*).
3. Quando João viu quão transtornada Maria estava com o que ele disse (*do físico ao mental*), sentiu-se muito arrependido (*do mental ao mental*), e decidiu pedir desculpas na mesma hora (*do mental ao físico*).
4. "Por toda sua vida ele tomou o silogismo que aprendera estudando a Lógica de Kiesewetter – 'Caio é um homem, homens são mortais, e portanto Caio é mortal' – como sendo verdadeiro apenas de Caio, não de si mesmo. Caio era um homem, um homem no sentido abstrato, e por isso o silogismo aplicava-se a ele: mas Ivan Ilytch não era Caio,

e não era um homem no sentido abstrato" (*do mental ao mental*).<sup>2</sup>

Esses exemplos exigem uma breve explicação e comentário.

(1) exemplifica o caso em que uma sensação corporal, um evento mental, causa uma instância de comportamento que não é uma ação, mas "uma reação, algo que é involuntariamente extraído de uma pessoa. A dor é a razão pela qual a pessoa grita, mas não a razão da pessoa, isto é, não é a razão que a pessoa tem ou adota para se comportar do jeito que se comporta. A noção de adotar uma razão e decidir agir baseado nela não se aplica a esse caso, em que não há nenhuma ação envolvida.

(2) e (4) exemplificam causação do mental ao mental, mas de maneiras diferentes. (2) é similar a (1). Assim como a dor torácica de uma pessoa suscita um grito involuntário, essa mesma dor momentaneamente produz pânico nessa pessoa. O pensamento de que ela poderia estar sofrendo um ataque cardíaco a acomete espontaneamente. Em (4), contudo, o movimento dos pensamentos expressos pelas frases "Caio é homem, homens são mortais" ao pensamento "Caio é mortal" é um passo lógico que Ivan Ilytch naturalmente atribui a Caio, mas não pode dizer que esse seja seu próprio caso devido à perspectiva de sua própria aniquilação. Quando pensamos logicamente tentamos ver a implicação de uma coleção de enunciados e nos permitimos guiar por isso. O movimento do pensamento ao pensamento exemplifica não apenas um processo causal, mas também um processo racional, e que em algum sentido está sob nosso controle.

O caso mais complexo é representado por (3). Um estado físico de coisas no mundo – Maria ficando atormentada, chorando, e assim por diante – ajuda a cau-

sar a percepção de João de que ela está atormentada (do físico ao mental). Essa percepção, por sua vez, faz com que ele se sinta arrependido, e este estado mental leva a um desejo de pedir desculpas (do mental ao mental). O desejo é apoiado por uma crença não enunciada de que pedir desculpas servirá para remediar a ofensa e trazer o ânimo de Maria de volta ao equilíbrio. Como resultado do desejo de consertar a situação e da crença de que pedir desculpas é a maneira correta de alcançar isso, João age correspondentemente (do mental ao físico).

Epifenomenalistas comprometidos devem negar que estados mentais possuam qualquer eficácia em todos esses exemplos. Eles concederão, naturalmente, que o físico origina o mental, mas não que o mental possa causar ocorrências físicas, ou mesmo desempenhar um papel no surgimento de outros estados mentais. Com isso em mente, examine o Exercício 7.1 antes de ir adiante.

### **EXERCÍCIO 7.1**

Qual é sua resposta à ameaça do epifenomenalismo? Considere sua reação e então leia minha resposta a seguir.

O epifenomenalismo é uma doutrina profundamente repulsiva por várias razões. Primeiramente, nossa experiência de primeira pessoa, como agentes, testemunha fortemente contra ela. Se eu atravessasse a rua para comprar alguns doces em uma loja, as razões para minha ação incluirão minhas crenças e desejos. Eu atravessei a rua porque queria uns doces e porque acreditava que poderia conseguilos na loja do outro lado. Se eu não tivesse desejado os doces, ou se acreditasse que lá não os encontraria, então não teria atravessado a rua (assumiremos que eu

não tenho outra razão para atravessar a rua). Isso certamente se parece com uma explicação causal do que eu fiz, chegando mesmo a implicar o condicional contrário-aos-fatos (que foi tomado como ponto alto da explicação causal), enunciando o que teria acontecido se, contrariamente ao que de fato aconteceu, eu não tivesse desejado os doces ou não tivesse acreditado que os encontraria na loja.

Em segundo lugar, estendemos a outras pessoas o que acreditamos ocorrer em nosso caso, a saber, que seus desejos e crenças também operam para explicar e causar seus comportamentos, que por sua vez produzem outras mudanças no mundo. De fato, basta olhar ao nosso redor para testemunhar a contribuição das mentes humanas às inumeráveis diferenças na ordem das coisas – sua criação e preservação, assim como sua destruição e deterioração. Mais uma vez, epifenomenalistas devem negar a verdade disso tudo. Para eles, nossas maneiras cotidianas de falar sobre a ação humana devem representar um enorme conjunto de erros, uma ilusão sistemática de que estados mentais fazem uma diferença, causalmente, para o modo como as coisas são.

Em terceiro lugar, uma peculiaridade de ulterior do epifenomenalismo é que se qualquer um acreditasse que ele fosse verdadeiro, e quisesse dizer isso, ele não seria capaz de fazê-lo, porque seus desejos e crenças seriam causalmente ineficientes. Paradoxalmente, qualquer um que acredite que o epifenomenalismo é verdadeiro, e consegue dizer isso, abandona a doutrina no próprio ato de afirmá-la. A asserção de uma crença pressupõe que ela própria, juntamente com um desejo de comunicá-la, tenha originado sua expressão verbal externa.

Compartilho da concepção de vários pensadores de que nossa própria experiência possa assegurar-nos de que o epi-

fenomenalismo seja mais falso do que verdadeiro, ainda que a especulação metafísica possa fazê-lo parecer verdadeiro. A melhor atitude em relação ao epifenomenalismo talvez seja aquela recomendada por Tyler Burge, que apropriadamente observou que ele pode ser visto como uma forma menor de ceticismo. Embora ninguém o leve a sério na prática, ainda assim é útil para aguçar nosso pensamento sobre a natureza das relações mentais e físicas.<sup>3</sup>

O epifenomenalismo coloca o desafio de como explicar da melhor maneira nossas usuais intuições de senso comum de que a mente efetivamente faz uma diferença ao modo como o mundo é. Compare o epifenomenalismo com o solipsismo. Ninguém seriamente adota o solipsismo ou o ceticismo do tipo que sustenta que não podemos saber que há um passado, ou que as leis da natureza continuarão operando no futuro assim como operam no presente. Mas o céptico, ao apresentar interrogações sobre se podemos conhecer aquilo que irrefletidamente acreditamos conhecer, desempenha importante papel em nos alertar para a necessidade de justificarmos nossas intuições.

### Razões para ações como causas

Como as ideias de Davidson sobre a explicação das ações supostamente levaram ao epifenomenalismo? Para explicar isso, preciso antes fornecer um esboço rápido da teoria de Davidson, segundo a qual razões para ações são causas. Em um artigo seminal, *Actions, Reasons and Causes*<sup>4</sup>, Davidson sustentou que as razões que as pessoas apresentam para explicar e justificar o que fazem funcionam

de modo padronizado como causas para essas ações.<sup>4</sup> À primeira vista, isso pode parecer incontrovertido. Afinal, se as razões de uma pessoa para suas ações não são as causas das ações que elas racionalizam, de que outro modo se poderia conceber sua relação? João desculpou-se com Maria. Por quê? Porque notou que feriu os sentimentos dela, desejo que esse não fosse o caso, e acreditou que dizer estar arrependido poderia reparar o dano. Como vimos anteriormente, isso certamente se parece com uma explicação causal, a ponto de implicar uma condicional contrafatual: caso João não tivesse desejado pedir desculpas, ou não tivesse acreditado que pedir desculpas seria bom, ou ambos, então ele não teria pedido desculpas. A ocorrência da ação de João é condicionada à existência da combinação apropriada de desejo e crença, e isso fornece uma forte razão para pensarmos que estamos lidando com uma explicação causal. Por que não deveríamos, então, na ausência de uma teoria mais plausível, aceitar que as razões que um agente dá para suas ações são as causas daquelas ações?

Uma razão foi fornecida pelo behaviorismo analítico, mantido vivo pelos elementos neobehavioristas na filosofia de Wittgenstein. Um grande problema com o behaviorismo é que, privando os estados mentais de um *status* interno por meio de sua redução a padrões de comportamento real e possível, eles não podem mais ser invocados para explicar o comportamento, incluindo as ações, uma vez que isso seria equivalente a tentar explicar um exemplo de conduta em termos dele mesmo. Contudo, parece claro que os exemplos (1)-(4) que forneci são todos instâncias genuínas de casos em que comportamento é suscitado por eventos mentais subjacentes.

\* N. de T.: Em português: *Ações, razões e causas*.

Um wittgensteiniano comprometido, tal como Norman Malcolm, que se voltou decididamente contra a concepção cartesiana dos estados mentais como sendo internos e privados, viu-se forçado a negar que desejos possam ser causas de ações.<sup>5</sup> Malcolm dá um exemplo de duas pessoas jogando xadrez. Um jogador pergunta ao outro: "Qual foi a causa para fazer aquele movimento?", e o oponente responde: "Não houve uma causa; eu quis fazer aquele movimento". Malcolm alega que essa resposta demonstra que o movimento do oponente não pode ser causado no sentido humeano padrão de causa. Antes de continuar, faça uma pausa para examinar no Exercício 7.2.

### **EXERCÍCIO 7.2**

Você concorda com a alegação de Malcolm? Considere sua própria resposta e depois continue lendo.

Malcolm está seguramente errado neste ponto. É verdade que "Qual foi a causa para fazer aquele movimento?" pode ser interpretado como "Quem, ou o que lhe obrigou, ou forçou, ou impeliu a fazer aquele movimento?", caso em que a resposta "Nada me obrigou ou forçou, foi minha própria decisão que me levou a fazer aquele movimento" é apropriada. Mas essa resposta apenas exclui certo tipo de causação do movimento, a saber, aquele em que alguma espécie de compulsão ou coerção atua sobre o agente. No exemplo dado, tal coerção é vista como emanando de uma fonte externa na figura de outra pessoa ou pessoas. Mas muito claramente o jogador de xadrez em questão poderia estar sofrendo de alguma espécie de compulsão psicológica ou impulso irresistível.

No entanto, o ponto central que precisa ser apresentado contra a alegação de

Malcolm é que, quando causas coercitivas ou incontroláveis estão ausentes, isso não significa que a ação do jogador de xadrez não tenha absolutamente *nenhuma* espécie de causa. O argumento de Malcolm não pode mostrar que quando a pessoa fez o movimento só porque quis, pensando que seria a melhor estratégia, seu desejo não desempenhou um papel causal na produção da ação. Não é de surpreender que Malcolm não desista de maneira tão fácil. Ele tem um estilo diferente de argumento que se funda em considerações sobre o que seria natural e fácil dizer a respeito de casos tais como esse que estamos examinando. Assim, Malcolm escreve:

Se podemos atribuir a B a causa de seu movimento, obtemos a seguinte consequência: não poderíamos dizer que o fato de B haver feito aquele movimento fosse "o efeito" de seu desejo de desconcertar seu oponente! Imagine B dizendo: "O fato de eu ter feito aquele movimento foi o efeito de meu desejo de desconcertar A!" Um exemplo diferente pode nos ajudar a ver este ponto: se eu quero escalar uma montanha para ver o que há lá em cima, e de fato a escalo por essa razão, não se pode dizer que minha escalada da montanha é "o efeito" de meu desejo de querer escalá-la para ver o que havia lá em cima – mesmo que este último possa ser chamado de "a causa" de minha escalada da montanha. Esses exemplos mostram que quando a causa de uma pessoa fazer X é sua razão, seu propósito para fazer X, então o fato de ela fazer X não é "o efeito" de sua razão ou propósito. Filósofos estiveram errados ao assumir que em todos os casos em que Y é a causa de X, X é o efeito de Y.<sup>6</sup>

O problema com esse estilo de argumento é que ele depende muito e injustificadamente da assim chamada "filosofia da linguagem ordinária". É verdade que as expressões que Malcolm cita possuem cer-

ta falta de naturalidade e um ar pomposo. No entanto, uma conclusão tão importante quanto a de que desejos e propósitos não têm a possibilidade de servir como causas em um sentido essencialmente familiar a Hume não pode se sustentar sobre uma base tão frágil. Como alternativa à expressão tediosa que Malcolm atribui a B, podemos imaginar B dizendo, por sua vez, "Eu fiz aquele movimento *porque eu quis desconcertar A*"; e é igualmente natural interpretar as palavras de B como fornecendo parte da causa de sua ação. A outra contribuição para a gênese da ação será dada, é claro, pela crença de B de que fazer o movimento que ele fez seria suficiente para perturbar A.

A segunda alegação de Malcolm é ainda menos satisfatória. Ele quer insistir que não se pode dizer de minha escalada da montanha que seja o efeito de meu desejo de escalá-la e ver o que há lá em cima, muito embora se possa dizer que meu desejo de escalá-la é a causa de minha escalada. De acordo com Malcolm, os filósofos têm estado equivocados ao pensar que quando Y é a causa de X, X pode ser sempre descrito como o efeito de Y. Essa afirmação soa claramente bizarra. Porém, presume-se que Malcolm quer reconstituir-la, reinterpretando-a de modo que signifique que X é a consequência não-causal de Y. Francamente, contudo, eu não consigo entender como X pode ser a consequência de Y, e ainda assim não ser causado por Y. Há uma forte aparéncia de autocontradição da parte de Malcolm neste caso.

A linguagem cotidiana talvez tenha a primeira palavra sobre tais questões, mas de modo algum a última. Jogando o jogo de Malcolm, por que não deveríamos dizer que o fato de alguém escalar uma montanha particularmente perigosa é a

consequência, o resultado, de um intenso desejo de provar que a natureza pode ser conquistada pelos esforços de um homem? Se superar a montanha foi o resultado da ambição desse homem, é certo, então, que a ambição ajudou a produzir a escalada da montanha? Se o homem não tivesse desejado escalar a montanha, se ele não tivesse tido a ambição de escalá-la, então ele não a teria escalado (eu deixo de lado a possibilidade de que alguém o tenha forçado a escalá-la). Isso se assemelha fortemente ao enunciado contrafactual "Se o ventilador não tivesse sido ligado na tomada, você não teria levado um choque". Ambos os enunciados conformam-se à segunda definição de causa apresentada por Hume em sua primeira *Investigation on Human Understanding*\* – "Se A não tivesse existido, B não teria existido" –, e são tão completamente causais quanto alguém possa querer. Além disso, assim como algo poderia intervir, impedindo o ventilador de lhe dar um choque elétrico – um rompimento no cabo de força, talvez –, algo também pode interromper a cadeia causal que vai do desejo de escalar a montanha à sua efetiva escalada – o aparecimento repentino de uma doença, talvez.

Malcolm invoca outro argumento bastante conhecido para tentar provar que desejos e intenções não podem ser causas: a alegação de que há uma conexão conceitual entre um desejo ou intenção e o objeto desse desejo ou intenção. Isso viola a exigência humeana de que causas e efeitos devem ser existências distintas e conectadas de maneira meramente contingente, não logicamente. Mas essa objeção é facilmente superada uma vez que os detalhes da explicação de Davidson sobre como razões explicam ações tenham sido dados, como veremos em breve.

\* N. de T.: Em português: *Investigação sobre o entendimento humano*.

Davidson chama de rationalização uma razão que explica por que um agente agiu como agiu. Uma rationalização tem dois componentes. Em primeiro lugar, ela justifica ou rationaliza uma ação ao mencionar um desejo ou pró-atitude (uma atitude favorável) em relação à ação descrita de certa maneira, e uma crença do agente de que a ação satisfaz aquela descrição. Como o próprio Davidson diz: "R é uma razão primária pela qual um agente desempenhou a ação A sob a descrição D, apenas se R consiste de uma pró-atitude em relação a ações com certa propriedade, e uma crença do agente de que A, sob a descrição D, possui aquela propriedade".<sup>7</sup> Assim, João pediu desculpas a Maria porque ele quis consertar as coisas, e acreditou que por meio de seu pedido de desculpas poderia fazer isso. Em segundo lugar, ao dizer que João agiu como agiu porque ele assim desejou e acreditou, a alegação de Davidson é que o "porquê" fornece as causas da ação de João, o desejo e a crença que, juntos, deram origem à ação.

Ao explicar um evento físico em termos de outro, quando nenhuma intervenção de agentes pessoais está envolvida, o aspecto causal da explicação é preservado, mas o elemento que o justifica não entra em cena. Quando explicamos por que uma chaleira ferveu, mencionando a chama que aqueceu a água, fornecemos a causa que faz com que a água ferva, mas não rationalizamos ou justificamos por que a água ferveu, uma vez que a chama não é um agente pessoal. Inversamente, o aspecto que justifica uma explicação pode ser preservado, mas o elemento causal pode estar faltando. Isso é o que acontece quando alguém inventa uma razão com o fim de justificar sua ação. Por exemplo, considere novamente o autoengano de *Lord Jim*, de Conrad (ver Capítulo 1, seção "Avaliando as doutrinas do acesso

privilegiado e da translucidez"). Jim repe-te a si mesmo que a razão de não resgatar dois homens que caíram do navio foi ter pensado que poderia ampliar seu conhecimento sobre os corações dos homens, não saindo em resgate. Ao dar a si mesmo essa razão para sua ação, seu objetivo, não reconhecido sequer por ele mesmo, foi aliviar sua incapacidade de agir, "tirar o corpo fora", como se costuma dizer. O que torna isto uma *mera* rationalização, e não a razão genuína pela qual ele não saiu em resgate, é que a explicação não menciona uma razão causalmente operante. A causa real da inação de Jim foi muito menos nobre: ele estava aterrorizado e não quis arriscar a própria pele, uma razão que ele não irá, ou não poderá, reconhecer para si próprio. E mesmo naqueles casos, em que tanto o elemento causal quanto o justificador estão presentes, na caracterização que Davidson faz de uma razão primária, está implícito, como vimos parcialmente, que a razão de João para pedir desculpas a Maria justificará sua ação apenas sob a descrição "pedir desculpas", e não sob qualquer outra descrição. Suponha que Maria interprete a ação de João como um exemplo de sarcasmo afiado, com a pretensão de ofendê-la. Então o desejo de João de consertar a sua situação com Maria, juntamente com sua crença de que um pedido de desculpas servirá para esse propósito, são causas de sua ofensa não-intencional. Mas elas não rationalizam ou justificam sua ação sob a descrição "ofender".

Um argumento contra a concepção de Davidson de que rationalizações são causas de ações se dá nas seguintes linhas: todos os enunciados causais, singulares e verdadeiros implicam leis causais; é apenas como espécies ou instâncias de certos tipos de evento que se pode dizer de um evento que é causa de outro. Dizendo de outra maneira, de acordo com Hume,

qualquer relação causal tem de ser vista como uma instância de uma regularidade geral na natureza. Assim, se um evento representante *a* do tipo de evento A causa um evento representante *b* do tipo de evento B, deve haver uma lei causal subjacente, descrevendo como eventos do tipo A tipicamente causam eventos do tipo B sob certas condições. Mas não há lei tal como "Sempre que pessoas como João querem consertar a situação com pessoas como Maria, e acreditam que podem fazer isso pedindo desculpas, elas dizem que estão arrependidas". Racionalizações de ações não podem ser refinadas até virarem enunciados de leis precisas com base nas quais predicações podem ser feitas.

Além disso, argumentou-se que leis devem ser verdadeiras como uma questão de fato, e não meramente em virtude dos termos que elas contêm. Isso quer dizer que enunciados que descrevem leis da natureza devem ser sinteticamente e não analiticamente verdadeiros. Mas rationalizações de ações parecem verdadeiras por definição e, por assim dizer, auxiliam a explicar o que nós queremos dizer com termos tais como "desejar" e "acreditar". Assim, se alguém deseja uma carteira de cigarros mais do que qualquer outra coisa, e acredita que pode conseguir uma na loja do outro lado da rua, então, a menos que ele mude de ideia no último minuto, ou morra em um acidente de carro, e assim por diante, ele vai atravessar a rua em direção à loja. Se ele não foi impedido de ir à loja e mesmo assim não foi até lá, nós concluiríamos que ou bem ele se livrou do desejo por cigarros, ou este foi suplantado por um desejo mais forte – não morrer de câncer de pulmão, por exemplo –, ou ele não acredita mais poder encontrar cigarros naquela loja. Nós chegamos a essa conclusão meramente refletindo no significado dos termos "desejar" e "acreditar". É parte integral do que significa dizer que

um agente que quer uma coisa X e acredita que ao fazer *a* pode obter X, que a menos que seja impedido, tenha desejos ainda mais fortes, mude de ideia e assim por diante, então desempenhará a ação *a*. Em outras palavras, parece existir alguma espécie de conexão lógica ou conceitual entre as rationalizações e as ações por elas explicadas que as impedem de serem instâncias de leis e, portanto, por sua vez, as impedem de serem causas de ações.

Antes de passar para a próxima seção, examine o Exercício 7.3.

### **EXERCÍCIO 7.3**

A alegação de que existe conexão conceitual entre ações e razões para ações que impedem razões de serem causas foi debatida durante muitos anos. Você consegue ver alguma maneira de resolver esse problema? Discuta suas ideias com outra pessoa antes de continuar com a leitura.

### **A resposta de Davidson: a teoria causal da ação intencional**

Davidson tem uma resposta que, se for bem-sucedida, pode contornar ambas as objeções esboçadas anteriormente. De acordo com Davidson, análises da causalidade devem proceder em dois níveis diferentes. No nível inferior, enunciados causais singulares descrevem uma relação causal que se dá entre dois eventos particulares. Essa relação causal se dá entre os eventos independentemente de como esses eventos são descritos. Por exemplo, considere a frase: "Riscar o fósforo no lado da caixa-de-fósforos causou que o fósforo se acendesse". O evento descrito como "riscar o fósforo" é causalmente relacionado com o evento descrito como

"o fósforo ardendo em chamas". São os próprios eventos que estão causalmente relacionados, e não as descrições dos eventos (ver Capítulo 6, seção "Searle e o problema mente/corpo"). Uma vez que a relação causal não se dá entre entidades linguísticas, mas entre os próprios eventos no mundo lá fora, diz-se das relações causais que elas ocorrem em um nível ontológico. Uma vez que a relação causal ocorre entre eventos e não entre descrições de eventos, ela ocorre, como eu disse um momento antes, independentemente de como os eventos são descritos. Ao contrário, diferentes descrições de eventos são diferentes em virtude de pinçarem diferentes propriedades desses eventos, e assim pode-se pensar em Davidson como mantendo que não é em virtude de como eventos são descritos, ou em virtude das propriedades pinçadas naquelas descrições, que eventos estão causalmente relacionados, mas meramente em virtude do fato de que um evento, *a*, está causalmente relacionado com outro evento, *b*.

No entanto, devemos distinguir um nível superior de causalidade, a saber, aquele em que as explicações causais são dadas. Explicações causais empregam o que veio a ser conhecido como o modelo nomológico-dedutivo de explicação, que prevalece nas ciências. Esse modelo consiste em um argumento dedutivo com duas premissas.

A primeira premissa consiste no enunciado de uma lei que descreve a relação de causa e efeito que se dá entre certas classes de eventos. Davidson enfatiza que tais leis são o que ele chama de leis estritas, com o que ele quer dizer duas coisas:

1. Leis são enunciados gerais verdadeiros que não contêm cláusulas do tipo "outras coisas sendo iguais". A regularidade que elas enunciam tem validade universal e sem qualificação. Com efeito, ele trata o

universo como um sistema físico fechado (ver Capítulo 3, seção "Uma breve retrospectiva histórica").

2. Elas embasam enunciados contrafatuais. Um enunciado contrafactual afirma que, se uma causa não tivesse ocorrido, então seu efeito também não deveria ocorrer. Isso quer dizer que um efeito não necessariamente sucede uma causa, mas que sua ocorrência depende daquela causa. Isso torna a lei tão determinista quanto a natureza pode ser. Em outras palavras, a noção da dependência do efeito em relação à causa auxilia a captar a ideia de que a causa é naturalmente necessária para o efeito ocorrer.

A segunda premissa consiste em um enunciado singular que descreve a ocorrência do evento que é a causa. Colocando-se esses dois enunciados conjuntamente em um argumento, podemos deduzir a ocorrência do efeito como conclusão do argumento. O modelo de explicação segue o seguinte padrão:

1. Se certas condições iniciais C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>... C<sub>n</sub> estão satisfeitas, então o evento E ocorrerá. (C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, etc. constituem as condições que precisam ser preenchidas para que o evento E ocorra.) Este é um enunciado da lei causal.
2. E, nesta instância, C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>... C<sub>n</sub> estão satisfeitas. (Enunciado singular asserindo que as condições iniciais descritas em (1) estão satisfeitas.)
3. Portanto, E ocorrerá.

Colocando isso de maneira menos abstrata, suponha que existe uma lei física que enuncia quanto um certo tipo de metal, digamos, cobre, expandirá se for aquecido de um grau centígrado. (Esta lei

enuncia o coeficiente de expansão linear do cobre.) No que toca a um específico pedaço de cobre que foi aquecido de um grau centígrado, podemos deduzir que ele terá expandido segundo a especificação da lei. O modelo de explicação resultará assim:

1. Se o cobre é aquecido de um grau centígrado, então expandirá por certo número de frações de um milímetro (lei causal).
2. Este pedaço de cobre foi aquecido de um grau centígrado.
3. Portanto, este pedaço de cobre terá sido expandido segundo a especificação da lei.

O que é deduzido não é o próprio evento, mas um enunciado (3) que descreve o evento, e é deduzido dos dois outros enunciados (1) e (2). É por isso que Davidson sustenta que a explicação causal se dá entre enunciados que descrevem eventos, e não entre os próprios eventos, uma vez que estes são entidades não-lingüísticas. Para uma explicação causal ser possível, tudo depende de como os eventos são descritos. Suponha, por exemplo, que minha caneta-tinteiro é feita de cobre. Então, embora seja verdade que minha caneta-tinteiro expandirá tantas frações de um milímetro se for aquecida de um grau centígrado, um enunciado que diga que minha caneta-tinteiro foi aquecida de um grau centígrado não pode ser substituído por (2) na explicação anterior para que a conclusão (3) ocorra logicamente. A explicação

1. Se o cobre for aquecido de um grau centígrado, então, expandirá tantas frações de um milímetro.
  2. Minha caneta-tinteiro foi aquecida de um grau centígrado
- e não pode gerar a conclusão.
3. Minha caneta-tinteiro expandiu de certa fração de um milímetro

porque a lei não foi formulada em termos do que acontece com uma caneta-tinteiro, mas apenas em termos do que ocorre com cobre. Assim, pode-se ver que, com respeito a explicações causais que empregam leis, a possibilidade de dar uma explicação depende de como os eventos são descritos. Em virtude de ser feita de cobre, minha caneta-tinteiro expandirá tantas frações de um milímetro se aquecida de um grau centígrado, e, sob a descrição de que é cobre, a dedução de que ela expandirá quando aquecida pode ser feita. Mas sob a descrição "o presente que eu recebi em meu aniversário de 21 anos", a dedução uma vez mais se torna impossível, mesmo que seja verdade que quando aquecido de um grau centígrado o presente que eu recebi em meu aniversário de 21 anos se expandirá de uma certa fração de um milímetro.

Davidson alega que todos os enunciados causais singulares deverão ser embasados por leis, mas que isso não significa que as descrições que aparecem no enunciado causal singular serão necessariamente as mesmas que aquelas que ocorrem no enunciado das leis. Sua alegação implica, contudo, que deve haver descrições dos eventos que aparecem em enunciados causais singulares que instanciem leis causais, mesmo que não estejamos cientes de quais sejam essas descrições. Isso é comumente o caso. Para usar um de seus próprios exemplos, ventos fortes podem ocasionar a queda de uma ponte, mas não há leis causais descrevendo relações entre ventos e pontes. Ainda assim, em algum nível haverá leis empregando conceitos que podem ser usados para caracterizar o movimento do ar e a queda de estruturas que servirão de base para explicação e predição da queda de pontes em certas ocasiões. Davidson apresenta o argumento de que enunciados causais singulares devem ser

embasados por leis causais da seguinte maneira:

Ele pode querer dizer que "A causou B" implica algumas leis particulares envolvendo os predicados usados nas descrições "A" e "B", ou ele pode querer dizer que "A causou B" implica que existe uma lei causal instanciada por algumas descrições verdadeiras de A e de B. Obviamente ambas as versões da doutrina de Hume dão sentido à alegação de que enunciados causais singulares implicam leis, e ambos sustentam a concepção de que explicações causais envolvem leis [...] Mas a segunda versão pode ser usada de modo a adequar-se à maioria das explicações causais; ela serve igualmente bem para racionalizações.<sup>8</sup>

Davidson pode usar a distinção entre enunciados causais singulares e explicações causais para evitar a segunda objeção – as razões de uma pessoa para suas ações não podem ser causas – da seguinte maneira. A segunda objeção você deve lembrar, dizia que o fato de que explicações de ações em termos dos desejos e crenças do agente acabavam por mostrar-se necessariamente verdadeiras – excluía aquelas explicações do âmbito das explicações causais, porque para que eventos possam ser relacionados como causa e efeito eles devem ter existências distintas, nenhuma das quais logicamente implica a existência da outra. A resposta de Davidson é salientar que o modo como um evento é descrito pode fazer com que ele pareça logicamente relacionado a outro evento e, portanto, não possa servir como uma causa genuína. O que precisamos nos dar conta é de que apenas descrições de eventos são logicamente relacionadas, não os próprios eventos. O fato de um evento poder ser descrito em termos que o conectem logicamente com seu efeito não pode mostrar que os eventos em questão não podem es-

tar relacionados como causa e efeito. Para ver isso, tome qualquer enunciado causal singular, tal como "Colocar a chaleira no fogo causou a fervura da chaleira". A ação descrita como "Colocar a chaleira no fogo" pode ser facilmente redescrita como "a causa da fervura da chaleira". Se essa redescricao é então recolocada no enunciado singular original, obtemos: "A causa da fervura da chaleira causou a fervura da chaleira". Mas embora esse enunciado seja analiticamente verdadeiro, isso não impede que colocar a chaleira no fogo seja a causa da fervura da chaleira, porque a relação lógica é obtida puramente em virtude de como os dois eventos são descritos. No nível dos eventos concebidos como particulares não-abstratos, eles ainda contam como existências distintas, e assim podem manter uma relação causal entre si.

Uma boa ilustração do mesmo ponto é fornecida pelo exemplo das substâncias descritas como venenosas. Descrever algo como um veneno é caracterizá-lo em termos de seus efeitos típicos – a doença e possível morte que ele ocasiona. É, assim, analítico dizer que, tudo permanecendo idêntico (p. ex., nenhum antídoto tendo sido administrado antes), um veneno causa morte e doença; se ele não ocasionasse tipicamente esse efeito, não seria um veneno. Mas venenos podem ser descritos em outros termos, o que mostra que sua relação para com os seus efeitos é puramente contingente afinal. Cianeto de sódio tipicamente causa doença ou morte se ingerido em quantidade significativa, e o enunciado de que cianeto de sódio causa doença ou morte não é necessariamente verdadeiro, e sim algo que acontece puramente como uma questão de fato.

Além disso, do mesmo modo que causas às vezes são descritas em termos de seus efeitos típicos, efeitos são às vezes

caracterizados em termos de suas causas. É analítico dizer que ovos fritos foram produzidos por meio de fritura, e que refeições cozidas resultaram de cozimento. Contudo, a fritura do ovo ou o cozimento da refeição foram causalmente responsáveis pelo estado frito ou cozido em que o ovo e a refeição, respectivamente, passaram a apresentar. Podem-se encontrar outras descrições de mudanças internas que resultam quando a comida é cozida, significando que sob tais descrições não há mais qualquer aparência de ligação lógica entre o cozimento da comida e seu estado final.

Aquelas descrições de desejos e intenções que especificam o que eles sejam – descrições individualizadoras – são como a descrição de algo como um veneno, na medida em que tais descrições implicam o que tipicamente ocorre como resultado de se ter o desejo ou a intenção em questão. Uma intenção de cortar a grama resultará, se tudo mais se mantiver idêntico, na grama sendo cortada. Isso de fato é o que significa ter uma intenção de cortar a grama, em vez de alguma outra intenção, e não desistir ou ser impedido por fatores intervenientes, e assim por diante.

Contudo, de acordo com Davidson, a intenção e a ação resultante, neste caso cortar a grama, deve ter descrições sob as quais elas possam aparentar instanciar uma lei estrita. Uma vez que Davidson pensa que as únicas leis estritas são leis físicas, isso significa que desejos, intenções e crenças que também rationalizam ações particulares devem todas possuir descrições físicas. Assim, Davidson escreve:

As leis cuja existência é requerida se razões são causas de ações não lidam, podemos estar seguros, com os conceitos com os quais rationalizações devem lidar. Se as causas de uma classe de even-

tos (ações) enquadram-se em uma certa classe (razões), e há uma lei que embase cada enunciado causal singular, não resulta haver uma lei conectando eventos classificados como razões a eventos classificados como ações – as classificações podem até mesmo ser neurológicas, químicas ou físicas.<sup>9</sup>

### Monismo anômalo :

Podemos agora examinar a posição de Davidson que ficou conhecida como monismo anômalo. Apresentarei um esboço da tese, enunciando as proposições que a constituem, comentando depois cada uma delas.

1. Eventos mentais causam eventos físicos.
2. Se dois eventos são causalmente relacionados, então deve haver uma lei física estrita sob a qual eles podem ser subsumidos. (Leis estritas não contêm nenhuma cláusula *ceteris paribus*, isto é, nenhuma cláusula “o resto mantendo-se idêntico”, e são parte de um sistema fechado, de modo que tudo que possa afetar os eventos dentro do sistema deve ser necessariamente incluído no sistema.) Dizendo de outra maneira, causa e efeito devem ter descrições sob as quais eles instanciam uma lei física estrita.
3. Não há leis psicofísicas estritas. Isto é, não há leis estritas conectando eventos mentais sob suas descrições mentais a eventos físicos sob suas descrições físicas. Isso, com efeito, se traduz em uma rejeição da teoria de identidade tipo/tipo que almeja a uma redução ontológica de eventos mentais a eventos físicos (ver Capítulo 1, seção “O caráter holístico dos es-

tados mentais e sua relação com o estado de consciência", e Capítulo 3, seção "Racionalidade e normatividade").

### Comentário

A primeira proposição esboçada na seção anterior representa a rejeição por parte de Davidson do epifenomenalismo. Eventos mentais podem funcionar, e de fato funcionam, como causas. Eles fazem uma diferença ao que acontece fisicamente no mundo: as razões de uma pessoa para suas ações são as causas daquelas ações.

A segunda proposição na seção anterior veio a ser conhecida na literatura em torno deste tópico como o princípio da causalidade nomológica. Davidson não argumenta em favor dessa posição, mas se contenta em indicar que ela pareceu verdadeira para muitos filósofos, dentre eles notoriamente Hume e Kant, entre outros. Dado que, de acordo com a primeira proposição, eventos mentais são causalmente relacionados a eventos físicos, deve haver uma lei estrita sob a qual tais eventos possam ser subsumidos, e esta deve ser uma lei física, de modo que deve haver uma descrição física do evento mental sob a qual ele pode aparecer na lei física estrita. Assim, eventos mentais devem ser eventos físicos. Esta é a tese do monismo de eventos. Ao contrário do que afirma o dualismo, não existem dois tipos de eventos, aqueles que são puramente físicos e aqueles que são puramente mentais. Em vez disso, todos os eventos são eventos físicos, mas alguns eventos físicos também possuem descrições mentais. De acordo com a terceira proposição na seção anterior, eventos mentais não podem ser reduzidos sistematicamente a eventos físicos, seja por redução analítica ou ontológica. Assim, não haverá leis causais psicofísicas estritas relacionando

o mental, descrito como mental, a eventos físicos.

A razão de Davidson para rejeitar a redução ontológica do mental ao físico é que as atitudes proposicionais – crenças, desejos, intenções, e assim por diante – são holísticas, e isso traz uma restrição normativa à atribuição dessas atitudes a outras pessoas. Deve, ele afirma, haver um alto grau de coerência entre as atitudes, pois do contrário estaríamos tentando imaginar um ser totalmente irracional, e isso seria equivalente a pensar em um animal sem pensamentos: "Na medida em que não logramos descobrir um padrão coerente e plausível nas atitudes e ações de outros, simplesmente abdicamos da chance de tratá-los como pessoas". Não há contrapartida desse ideal constitutivo da racionalidade na teoria física, e portanto não pode haver a redução sistemática, indiscriminada, do mental ao físico que a teoria da identidade tipo-tipo requer.

É na impossibilidade desse tipo de redução que consiste a natureza anômala do mental, seu caráter não-legal. Uma vez que não pode haver leis psicofísicas, eventos mentais não podem ser nem explicados ou previstos em termos puramente físicos. Alguém poderia saber a história física completa do mundo, e mesmo assim ser incapaz de explicar ou prever um único evento mental, muito embora todo evento mental necessariamente se enquadrasse sob uma ou outra descrição física. A rejeição de Davidson da redução do mental ao físico em uma base tipo-tipo, juntamente com sua insistência em que eventos mentais devem possuir descrições físicas, implica que ele sustenta alguma versão da identidade espécime-espécime do mental com o físico. Com o objetivo de explicar a relação entre as características mentais dos eventos físicos e as características físicas desses eventos,

Davidson desposa a superveniência. Dizer que o mental sobrevém ao físico é fazer a alegação modal de que não pode haver eventos que sejam similares em todos os aspectos físicos, mas difiram em algum aspecto mental (ver Capítulo 6, seção "A doutrina da superveniência"). Assim, se o estado mental de uma pessoa muda de M1 a M2, isso quer dizer que deve haver alguma mudança nos estados físicos sobre os quais M1 e M2 sobrevêm. Do mesmo modo, os estados mentais de duas pessoas não podem ser diferentes, de acordo com a tese da superveniência, a menos que haja uma diferença nas bases físicas subvenientes que fundamentam aqueles estados mentais.

### Objeções ao monismo anômalo

A primeira objeção que devemos considerar deriva de Jaegwon Kim, que diz que a posição de Davidson o compromete com a concepção de que "eventos são causas ou efeitos apenas na medida em que eles instanciam leis físicas".<sup>10</sup> Mas isso, defende Kim, é o mesmo que dizer que estados mentais são causalmente eficientes apenas em virtude de suas características físicas, e não das mentais. O mental não causa nada mais *qua* mental; não é *como* mental que ele possui um papel causal para desempenhar, mas apenas devido a suas propriedades físicas, sob as quais apenas ele pode ser visto como instanciando uma lei física estrita, que todos os enunciados causais singulares verdadeiros devem instanciar. Portanto, o monismo anômalo leva diretamente ao epifenomenalismo. A objeção de Kim é complicada. Consequentemente, antes de prosseguir a leitura, faça uma pausa, e examine o Exercício 7.4.

### EXERCÍCIO 7.4

Kim está certo ao alegar que Davidson está comprometido a defender a tese de que propriedades mentais de eventos físicos são causalmente impotentes e, por conseguinte, que a mente não faz diferença causal alguma no mundo?

Davidson despreza a objeção de Kim. Relações causais para ele se dão entre eventos, particulares não-abstratos, independentemente de como eles são descritos. Ele não está comprometido com a concepção de que apenas em virtude de suas propriedades um evento possa ser a causa de outro: "Dada essa concepção extensionista das relações causais, não faz sentido literal algum [...] falar de um evento causando algo como mental, ou em virtude de suas propriedades mentais, ou como descrito de uma ou outra maneira".<sup>11</sup> Ele também diz de outra maneira: "Eventos instanciam uma lei apenas como descritos de um modo em vez de outro, mas não podemos dizer que um evento causou o outro apenas como descrito".<sup>12</sup> O fato de que todos os eventos mentais possuem descrições físicas faz com que sejam plenamente partes do mundo físico, na teoria de Davidson, mas seria errado insistir que eles só podem funcionar como causas na medida em que possuem propriedades físicas. Quais são as propriedades que os eventos possam ou não ter é irrelevante para seus poderes causais, no que toca a Davidson. Desse modo, ele se esquiva da questão sobre se eventos mentais causam eventos físicos corporais em virtude de suas propriedades mentais ou em virtude de suas propriedades físicas, rejeitando o dilema que lhe é imputado. Para ele, não há dilema a ser confrontado. Eventos mentais não são meros epifenômenos, mas isso independe de quais ca-

racterísticas, mentais ou físicas, tais eventos possuem.

Isso, no entanto, é fortemente contra-intuitivo. Considere um evento que consiste em uma explosão que mata várias pessoas. Talvez, quando a explosão ocorrer, emita um clarão vermelho, em vez de esverdeado. Nós diríamos que a mera cor da explosão, considerada em si mesma, é uma característica ou propriedade da explosão que é causalmente irrelevante para seus efeitos devastadores, que dependem de outras propriedades mais pertinentes, tais como as forças geradas quando as substâncias químicas detonam, as temperaturas alcançadas pelos gases explosivos, e assim por diante. Em outras palavras, parece razoável concluir que algumas propriedades dos eventos, mas não outras, desempenham um papel causal em suscitar outros eventos específicos. A suspeita que Kim expressou a respeito do papel do mental no monismo anômalo não pode, portanto, ser inteiramente descartada. Proponho, na base do que Davidson disse até o momento, que deve permanecer alguma dúvida residual sobre o papel do mental se, para que figure em uma lei causal, deva atrair uma descrição física.

Um segundo receio é que, se eventos mentais são causados por eventos físicos subjacentes e têm existência e traços devidos inteiramente a tais eventos, então é difícil entender como eventos mentais possam ter poderes separados e independentes que lhes sejam próprios. Se  $M$  é o evento mental que sobrevém a um evento físico  $F$ , é questionável que poderes causais possam emergir em níveis superiores. "Não há", escreve Kim, "novos poderes causais que magicamente floresçam em  $M$  [o estado mental] acima e além dos poderes causais de  $F$  [o estado físico ao qual  $M$  sobrevém]. Nenhum novo poder causal emerge em níveis superiores".<sup>13</sup>

Searle está alerta para essa dificuldade, mas, com o sangue-frio que lhe é característico, argumenta que ela pode ser facilmente superada:

O fato de que características mentais são supervenientes em relação a características neuronais de forma alguma diminui sua eficácia causal. A solidez do pistão é superveniente em relação à sua estrutura molecular, mas isso não torna a solidez epifenomênica; e, similarmente, a superveniência causal de minha dor nas costas presente sobre microeventos em meu cérebro não torna a dor epifenomênica.<sup>14</sup>

Você deveria parar para analisar o Exercício 7.5 antes de prosseguir a leitura.

### **EXERCÍCIO 7.5**

Você concorda com Searle? Retorne ao Capítulo 6, seção "Searle e o problema mente/corpo", para refrescar sua memória a respeito dessa questão.

Infelizmente para Searle, sua proposta não funcionará. A superveniência da solidez do pistão relativamente à sua microestrutura não possui paralelo na superveniência da dor relativamente a microeventos cerebrais. A razão remete a algo que eu indiquei em minha discussão de Searle no último capítulo: diferentemente da liquidez da água, a solidez do pistão não é um fenômeno distinto e separado de sua microestrutura. A solidez do pistão é constituída por sua microestrutura. É aquela mesma estrutura vista de uma perspectiva macroscópica. No caso da solidez do pistão, a teoria da identidade é verdadeira. A solidez é a microestrutura, assim como a temperatura de um gás é a energia cinética média de suas moléculas. A razão pela qual a solidez do pistão

não é epifenomênica é que a sua solidez/microestrutura é, sem nenhuma dúvida ou ambiguidade, parte constitutiva do mundo físico. A solidez simplesmente é a microestrutura física, e, por conseguinte, não é surpreendente que ela não possa ser classificada como epifenomênica (ver Capítulo 6, seção "Searle e o problema mente/corpo").

Por outro lado, o fenômeno da consciência, como o próprio Searle admite, é um fenômeno separado e distinto dos microeventos físicos que estão em sua origem. Aqueles microeventos são partes de um reino objetivo, de terceira pessoa, público e espacial, o mundo físico. A consciência, por outro lado, é um fenômeno subjetivo, de primeira pessoa, privado. Aplicando a lei de Leibniz (ver Capítulo 2, seção "O argumento da dúvida"), somos forçados à conclusão de que a consciência não pode ser idêntica a eventos microfísicos, mas deve ser vista como uma espécie separada e distinta de ocorrência.

Uma dificuldade intimamente relacionada e bastante digna de consideração é colocada pela tese de que o mundo físico é causalmente fechado, e que causas mentais e físicas não determinam excessivamente seus efeitos. Contudo, para esclarecer o foco dessa objeção, alguma explicação será requerida (ver Figura 7.1). Um evento físico, F, um evento cerebral de alguma espécie, é causalmente suficiente para originar outro evento cerebral, F\*. Em outras palavras, o mundo

físico é causalmente fechado: todos os eventos físicos possuem causas físicas que são suficientes para suscitar tais eventos. Isso, como salientamos, constitui a base central do monismo anômalo. Ao mesmo tempo, um evento mental, M (por exemplo, uma sensação de dor), sobrevém a F. F determina a existência e o caráter de M, e M não pode mudar a menos que F mude. Para Searle, como vimos, a relação de superveniência é apenas causalidade: F causa M. Kim discorda dessa caracterização da superveniência, mas neste contexto o desacordo entre Searle e Kim não é relevante. M\* sobrevém a F\* que determina a existência e o caráter de M\* e o torna igual a um desejo de se livrar de M, a dor.

Na vida cotidiana poderíamos explicar a dor, M, minimamente como parte da razão pela qual M\*, o desejo de se livrar da dor, vem a existir. Em outras palavras, M desempenha um papel causal na produção de M\*. Contudo, parece claro a partir da Figura 7.1 que a razão real pela qual M\* existe não é por causa de M, mas apenas por causa de F e F\*. F atua para produzir F\*, e F\*, por sua vez, é a causa superveniente de M\*. M\* segue-se a M, mas apenas porque no nível básico subveniente F produz F\*. M não desempenha papel causal algum na geração de M\* – ele apenas aparece fazê-lo. Todo o trabalho causal real é feito no nível físico subveniente. M e M\* são meros epifenômenos adjacentes de F e F\* e que

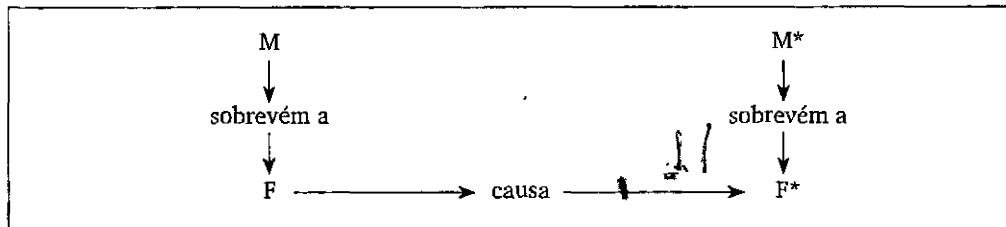


Figura 7.1 Superveniência mental/física e causação física.

não trazem nenhuma diferença causal ao mundo.

Mas, pode-se cogitar, mesmo se  $F$  é causalmente suficiente para suscitar  $F^*$ , por que  $M$  não deveria também ser parte da causa de  $F^*$ , e desse modo desempenhar um papel causal indireto na geração de  $M^*$ , e consequentemente evitar ser relegado a mero epifenômeno? Por que a Figura 7.1 não poderia ser modificada da maneira mostrada na Figura 7.2? A alegação que está sendo feita agora é que  $F^*$  é sobredeterminado por  $F$  e  $M$ . Isso quer dizer que ambos,  $F$  e  $M$ , operam como causas independentes e separadas para suscitar  $F^*$ .

Após exame cuidadoso dessas afirmações, e das Figuras 7.1 e 7.2, examine o Exercício 7.6 antes de prosseguir a leitura.

### EXERCÍCIO 7.6

Qual é sua reação? Pode um evento tal como  $F^*$  ser suscitado por duas causas independentes e separadas?

Filósofos têm sido muito cautelosos em aceitar que um único evento possa ter duas causas independentes, ambas sendo independentemente suficientes para produzir aquele evento. A intuição deles é de que, dados um evento  $C$  e duas causas postuladas  $A$  e  $B$ , há apenas três possibilidades disponíveis que sejam genuínas.

Primeiro,  $A$ , mas não  $B$ , foi a causa de  $C$ . Segundo,  $B$ , mas não  $A$ , foi a causa de  $C$ . Terceiro, nem  $A$  nem  $B$  foram a causa de  $C$ . (Para os propósitos desta discussão, esta última possibilidade pode ser desconsiderada.)

Para começar a entender por que a sobredeterminação causal tem sido rejeitada, considere o seguinte exemplo. Um homem está prestes a ser queimado em uma fogueira. Desejando assegurar-se de que ele não sofra, um de seus amigos joga um pequeno comprimido a seus pés. Esse comprimido, quando aceso, liberará uma fumaça altamente venenosa que vai matar a pessoa que está prestes a ser queimada imediatamente após as chamas alcançarem seus pés. Se o comprimido não funcionar por alguma razão, então o homem queimaré até a morte. Alternativamente, se o comprimido efetivamente funcionar, o homem morrerá muito antes de as chamas produzirem qualquer efeito significativo nele. Tanto o comprimido quanto a chama, considerados singularmente e separadamente, são causalmente suficientes para matar o homem. Mas agora suponha que alguém alegue que a morte do homem foi resultado de ambos, as chamas e o veneno trabalhando conjuntamente. Aqui é que gostaríamos de protestar: ele não pode ter sido morto por ambos os agentes. Deve ter sido um ou o outro. Se foi o veneno, então as chamas podem ser excluídas como sendo a causa. Do mesmo modo, se foram as chamas,

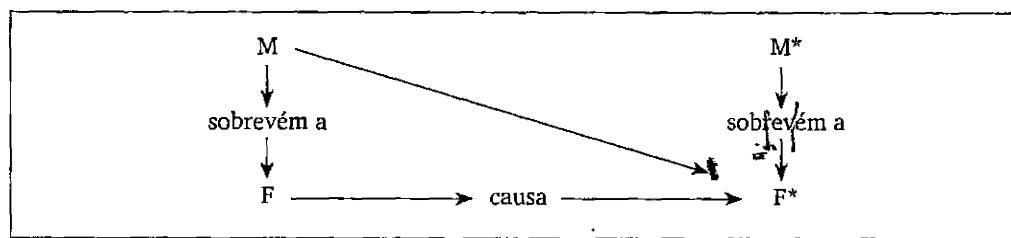


Figura 7.2 Determinação exacerbada?

então o veneno é excluído como a razão para a morte do homem.

Uma demonstração formal do argumento pode ser facilmente fornecida se lembrarmos que qualquer enunciado causal verdadeiro deve implicar um condicional contrário-aos-fatos. Em outras palavras, se X é a causa de Y, isso significa que, se X não tivesse ocorrido, então Y não ocorreria. Apliquemos isso ao caso em pauta e vejamos qual é o resultado:

Se as chamas são a causa da morte do homem, então sem a ocorrência das chamas o homem não teria morrido.

Contudo, no caso em que o veneno opere, o homem efetivamente morre no final; assim, o condicional contrário-aos-fatos, enunciando o que teria acontecido caso as chamas não tivessem ocorrido, é falso.

Por outro lado, considere:

Se o veneno é a causa da morte do homem, então sem o funcionamento correto do veneno o homem não teria morrido.

Contudo, se as chamas queimarem o homem, então ele morre, afinal de contas, e assim o condicional contrário-aos-fatos, enunciando o que teria acontecido caso o veneno não tivesse funcionado, é falso. As assertivas (1) e (2) não podem ser verdadeiras simultaneamente. Se (1) é verdadeira, então (2) deve ser falsa, e se (2) é verdadeira, então (1) deve ser falsa. Em outras palavras, (1) e (2) não podem operar como causas independentes e separadas. A morte do homem não pode ser exacerbadamente determinada: ela deve ou ter sido causada pelas chamas ou pelo envenenamento, mas não por ambos.<sup>15</sup>

Aplicando isso ao problema da causalidade psicofísica, se F de fato é causalmente suficiente para suscitar F\*, então, se a sobreDeterminação for rejeitada, simplesmente não há espaço para que

M possa operar. O problema poderia ser solucionado se M pudesse ser ontologicamente reduzido a F, isto é, se a dor pudesse ser identificada com eventos neurais no cérebro da pessoa. Mas então o que sobra da experiência de dor, e da consciência em geral, se esses fenômenos mentais afinal não são nada mais do que eventos físicos e químicos na cabeça? Reduzir fenômenos mentais a estados puramente físicos certamente impede aqueles fenômenos de serem causalmente inativos. Isso os traria completamente para dentro do mundo físico e os tornaria inteiramente parte de um sistema físico causal fechado. Mas isso só pode ser alcançado se negarmos o fato distintivo da consciência, o que é uma consequência inaceitável. Certamente, não é uma opção que o próprio Searle possa adotar à luz de sua rejeição do reducionismo materialista.

Outra opção seria abolir a relação de superveniência entre o mental e o físico, de modo que o mental não dependa do físico para sua existência ou caráter. Mas isso poderia parecer uma escolha ainda menos atraente, pois equivaleria a admitir que pudessem existir eventos puramente mentais, não-físicos, que pudessem permanecer exatamente como são, em total independência do mundo físico. Em outras palavras, somos levados diretamente de volta aos braços do dualismo substancial com todas as dificuldades que o acompanham.

Se o mental sobrevém ao físico, o que parece ser uma crença muito razoável, e se isso é feito para excluir a redução materialista, então temo que, pelas razões dadas, eu não possa ver como estados mentais possam não se tornar epifenômenos. Minha própria experiência, por outro lado, me diz que isso não pode ser o caso, e que meus próprios estados mentais, assim como os de outrem,

fazem uma diferença causal clara para o mundo. Como essas duas posições devem ser reconciliadas eu ainda não posso ver. Posso apenas deixar o problema em suas mãos.

### **Questões para reflexão**

1. As razões para ações podem ser as causas daquelas ações, como alega Davidson?
2. Todos os enunciados causais singulares devem ser embasados por leis causais?
3. Davidson nega que esteja comprometido com a alegação de que é apenas em virtude de suas propriedades físicas que um evento causa outro, mas ele está correto?
4. Se as propriedades mentais de eventos são epifenômenos, então como podem as razões de um agente justificar e explicar suas ações?
5. Pode-se provar que o epifenomenalismo é falso?
6. Se estados mentais são causados por estados do cérebro, pode haver causação genuína de cima para baixo, do mental ao físico?

### **Sugestões para leituras aprofundadas**

- MOYA, C.M. *The Philosophy of Action: An Introduction*. Cambridge, Polity, 1990 é uma excelente introdução às questões sobre a natureza da ação, sobre a vontade, sobre se as razões para ações podem ser causas, sobre monismo anômalo e sobre as contribuições de Davidson a essa área da filosofia. Altamente recomendado por sua clareza e abrangência.

- DAVIDSON, D. *Essays on Actions and Events*. Oxford, Clarendon Press, 1980 é um clássico do gênero, mas o leitor não familiarizado com Davidson deverá estar preparado para trabalhar duro.
- KIM, Jaegwon. *Mind in a Physical World*. Cambridge, MA, MIT Press, 1998 e \_\_\_\_\_. *Philosophy of Mind*. Boulder, CO, Western Press, 1996. Ambos os livros contêm discussões úteis sobre causação mental.
- HEIL, John e MELE, Alfred, eds. *Mental Causation*. Oxford, Clarendon Press, 1995 é uma coleção útil e atualizada, mas difícil.
- EVNINE, S. Donald Davidson. Cambridge, Polity, 1991 é uma introdução muito clara ao pensamento de Davidson.
- THALBERG, I. *Perception, Emotion and Action*. Oxford, Blackwell, 1977 contém boas discussões sobre a tese de Davidson de que razões para ações são causas e sobre o problema colocado por cadeias causais instáveis.

**NOTAS**

1. DAVIDSON, D. "Mental Events", in: \_\_\_, ed. *Essays on Actions and Events*. Oxford, Clarendon Press, 1980. Este artigo apareceu pela primeira vez em FOSTER, L. e SWANSON, J. W., eds. *Experience and Theory*. Cambridge, MA, University of Massachusetts Press e London: Dickworth, 1970.
2. TOLSTOY, L. "The Death of Ivan Ilyitch", in: *The Works of Leo Tolstoy*, vol. 15, Oxford, OUP, 1934.
3. BURGE, T. "Mind-Body Causation and Explanation", in: HEIL, John e MELE, Alfred, eds. *Mental Causation*. Oxford, Clarendon Press, 1995, p. 102-3.
4. DAVIDSON, D. "Actions, Reasons and Causes", in: \_\_\_, ed. *Essays on Actions and Events*.
5. ARMSTRONG, D. M. e MALCOLM, N. *Consciousness and Causality*. Oxford, Blackwell, 1984, p. 71.
6. Ibid., p. 72.
7. DAVIDSON, D. "Actions, Reasons and Causes", p. 4.
8. Ibid., p. 16-7.
9. Ibid., p. 17.
10. DAVIDSON, D. "Thinking Causes", in: HEIL, John e MELE, Alfred, eds, *Mental Causation*, p. 13.
11. Ibid.
12. Ibid., p. 6.
13. KIM, J. *Philosophy of Mind*. Boulder, CO, Western Press, 1996, p. 232.
14. SEARLE, J. *The Rediscovery of the Mind*. Cambridge, MA, MIT Press, 1992, p. 126.
15. Compare com a discussão de Samuel Guttemplan sobre este ponto em seu artigo "An Essay on the Mind", in: GUTTEPLAN, S. ed. *A Companion to the Philosophy of Mind*. Oxford, Blackwell, 1994, seção 3.4, p. 84-6.

# 8

## UMA PERSPECTIVA ARISTOTÉLICA

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender como Aristóteles evita os erros do dualismo e da identidade mente/cérebro atribuindo estados mentais e físicos a um único sujeito, o ser vivo.
- Entender a distinção feita por Aristóteles entre matéria e forma e a relação existente entre elas.
- Entender o que Aristóteles quer dizer com *psuche* e as diferentes espécies de *psuche* que as coisas vivas possuem.
- Entender por que a teoria de Aristóteles aventa a possibilidade de pessoas não-humanas.
- Entender a distinção entre uma capacidade, seu exercício e seu veículo.
- Entender as contribuições diferentes e complementares que os filósofos e os cientistas possam fazer à filosofia da mente.
- Entender e avaliar a doutrina do intelecto ativo incorpóreo.
- Entender como a teoria de Aristóteles evita a objeção das muitas mentes, apresentada pelo dualismo.
- Entender como a teoria de Aristóteles pode evitar o problema da interação mente/corpo.
- Entender o que é, na filosofia da mente, a falácia mereológica, e como a teoria de Aristóteles nos ajuda a evitá-la.

### Introdução

Assim como nos últimos anos tem ocorrido uma renovação do interesse a respeito da teoria ética de Aristóteles, e em particular sobre sua concepção de

virtude e vida eudaimônica, igualmente muitos filósofos têm-se voltado à filosofia da mente de Aristóteles a fim de obter valiosos insights sobre o problema mente/corpo. Vistos de uma perspectiva aristotélica, o dualismo de substâncias e a teoria



da identidade mente/cérebro compartilham um defeito comum: que eles localizam os estados mentais no local errado, o primeiro atribuindo estados da mente a uma alma incorpórea invisível, e a última, ao cérebro humano vivo. De acordo com Aristóteles, no entanto, o verdadeiro local dos estados mentais é o usual ser humano de carne e osso. Estados mentais não se unem apenas a mentes fantasmagóricas, e aspectos físicos, com igual exclusividade, a corpos corpóreos. Nem deve o cérebro, ou mesmo todo o corpo de uma pessoa, ser pensado como o possuidor de estados mentais. Pois não é o cérebro que sente dor ou pensa, nem mesmo o meu corpo, mas eu, a pessoa viva que unifica tanto atributos físicos quanto mentais em um único sujeito. Uma expressão moderna desta proposição é formulada por P. F. Strawson em seu livro *Individuals*:\*

O conceito de uma pessoa não deve ser analisado como aquele de um corpo animado ou de uma alma incorporada [...] o conceito de uma pessoa deve ser entendido como o conceito de um tipo de entidade, tal que tanto os predicados atribuindo estados da consciência quanto os predicados atribuindo características corpóreas, uma situação física, etc., são igualmente aplicáveis a um indivíduo deste tipo.<sup>1</sup>

A esse respeito Aristóteles tem uma grande afinidade com Wittgenstein, que escreveu, apesar de fazê-lo por razões inteiramente diferentes e relacionadas a como termos mentais adquirem seus significados: "Somente de um ser humano vivo e o que a ele se assemelha (comporta-se como ele) pode-se dizer: ele tem sensações; vê; é cego; escuta; é surdo; é consciente ou inconsciente".<sup>2</sup>

Aristóteles também pode ser percebido como assentando os fundamentos de um programa de biopsicologia teórica

científica. Então, tanto a filosofia quanto a ciência podem ser levadas a se sustentar na natureza da pessoa e de suas capacidades e estados psicológicos. Por outro lado, os materialistas redutivos, tais como os teóricos da identidade mente/cérebro e fisicalistas em geral, não dão atenção à psicologia em favor do físico; os dualistas, por insistir na incorporeidade da mente, e na não-fisicalidade dos estados mentais, põem estes para sempre além do alcance de investigação e explicação empíricas.

### Matéria e forma

A fim de avaliar integralmente a contribuição de Aristóteles à filosofia da mente, precisamos primeiro entender sua distinção entre matéria e forma, e em particular o que ele chama de "forma substancial". O termo "forma" no inglês contemporâneo sugere molde, mas essa não foi a intenção de Aristóteles. A forma substancial de uma coisa é, de maneira aproximada, aqueles traços que fazem com que algo seja o tipo de coisa que ela é, e esses traços, por sua vez, devem ser pensados como aqueles poderes ou capacidades que uma coisa desta espécie caracteristicamente tem e que portanto capacitam-na a realizar sua função. Um exemplo de Aristóteles que ilustra a distinção é aquele de um machado: a forma de um machado é sua capacidade de cortar madeira, e ele possui este poder em virtude do fato de que ele tem uma cabeça, com uma extremidade afiada, e um cabo no qual esta é firmemente fixada.<sup>1</sup> A matéria do machado é a coisa – a madeira e o metal – da qual ele é composto. A relação entre forma e matéria é lógica: a matéria e a forma não são em si mesmas substâncias, coisas que podem compor as partes constituintes de algo assim como as peças de um quebra-cabeça o compõem, mas que podem ser

\*N. de T.: Em português: *Indivíduos*.

desconectadas uma da outra e espalhadas, existindo independentemente uma da outra. Antes, forma e matéria são inseparáveis no sentido de que a forma não pode existir separada da matéria que ela qualifica, assim como a matéria precisa tomar uma forma ou outra. Em outras palavras, a forma é necessariamente materializada, e a matéria necessariamente enformada.

Um exemplo contemporâneo do que equivale à distinção forma/materia é fornecido por um balão. Um balão na verdade apenas é considerado um balão se ele possui a capacidade de conter o ar ou um gás de alguma espécie, tal como o hidrogênio ou o hélio. Não importará muito do que exatamente o balão é feito, assim como o quanto ele pode ser soprado e expandido pela introdução de ar ou de um gás. Agora deve ser lembrado que em *Urso Puff*, Leitão quer dar a Bisonho um balão por causa de seu aniversário para alegrá-lo, mas infelizmente, enquanto está correndo para a festa de Bisonho, Leitão tropeça e cai, e o balão estoura. Então ele não é mais um balão, mas apenas um pequeno, triste, úmido e flácido farrapo de borracha. Assim como Bisonho diz para Leitão: "De que cor era este balão quando ele – quando ele **era** um balão?" Sua qualidade de ser balão, sua forma substancial, literalmente sumiu com um estouro, e, sendo os balões o que eles são, isto é efetivamente não-restaurável.

### **EXERCÍCIO 8.1**

Forneça três exemplos pessoais para ilustrar o que Aristóteles quer dizer por forma substancial.

### **A relação da matéria à forma**

O termo "matéria" não deve ser entendido de modo estritamente literal – isto é,

como significando estreitamente apenas algumas coisas materiais, tais como terra ou metal, embora o conceito enquadre a ambas, juntamente com muitas outras. Os conceitos de matéria e forma são conceitos relativos: nada é intrinsecamente matéria ou forma, mas pode contar como uma ou outra quando em relação a alguma outra coisa. Por exemplo, um tijolo é uma forma em relação à argila da qual é composto, mas o tijolo, por sua vez, pode ser visto como compreendendo parte da matéria de uma parede. Por sua vez, uma parede é a forma que uma coleção de tijolos assume quando organizados de determinada maneira. Uma pilha de tijolos espalhados no chão de maneira aleatória, que então tenha cimento espalhado sobre ela para fixar os tijolos juntos, não é parede; os tijolos precisam estar nos locais certos a fim de servir para as funções certas. As paredes em relação às casas compõem sua matéria, e casas construídas com certa proximidade umas das outras são parte da matéria de vilas e cidades. De maneira semelhante, seres humanos racionais, por meio de sua interação, associação e combinação, são a matéria que compõe a sociedade civil e o corpo político.

A matéria da qual os seres humanos são compostos é, com certeza, extremamente complexa, não apenas porque um corpo humano vivo contém diferentes espécies de elementos naturais e os complicados compostos formados a partir deles, mas em sua estrutura organizacional. Você pode ter ouvido isso popularmente querendo dizer que um ser humano, em última análise, nada mais é do que certo número de galões de água, uns tantos baldes de cal, tanto mais de ferro, magnésio, fósforo e assim por diante, e isto é tudo que uma pessoa é. Isso não pode ser correto: uma pessoa não é somente um amontoado de materiais jogados no solo tanto quanto um muro não é um amontoa-

do de tijolos empilhados aleatoriamente. Meu computador não é apenas um grupo de fios, plástico e silício amarrados juntos de qualquer maneira. O que possui uma importância crucial é como todas essas espécies de coisas são organizadas. Seres humanos, assim como outros seres vivos, possuem uma estrutura complexa, e isso significa para Aristóteles que eles precisam ter órgãos que exerçam certas funções. Obviamente, eles podem fazer isso apenas em relação a outros órgãos que vêm juntos em um todo sistemático interconectado para constituir o organismo humano.

A matéria não pode existir sem ter alguma forma substancial: não pode haver matéria simples, sem traços característicos. Após o balão de Leitão ter estourado, o que restou ainda possui algumas capacidades; ele pode ser esticado, por exemplo, e soprado para formar vários balões pequenos. Ele poderia conceitivamente ser usado, junto com uma fita elástica, para selar a abertura de um jarro, e tenho vagas lembranças de uma juventude esbanjada que indicam que alguém pode fazer barulhos agudos e altos com a borracha de velhos balões falecidos.

Além das características substanciais de uma coisa, há formas accidentais que não são essenciais para fazer com que isto seja o tipo de coisa que ela distintamente é. Então o cabo de madeira de um machado pode ser substituído por um de plástico; na medida em que a capacidade do machado de cortar é preservada, a identidade do machado ao longo do tempo é preservada. Mas se o aço do qual a cabeça é composta fosse substituído, digamos, por borracha dura, o machado, estritamente falando, não seria mais um machado, pois ele não mais possuiria a capacidade de cortar: um machado com a cabeça embrorrachada é um machado apenas no nome. Do mesmo modo, quando o vinho se torna vinagre, e o leite, quei-

jo, não restam mais nem vinho nem leite, tendo sido perdida a identidade da substância original.

### Aristóteles e a *psuche*

Artefatos, como machados e balões, e comidas e bebidas, como queijo e vinho, são, é claro, inanimados. Mas coisas vivas, de acordo com Aristóteles, não menos do que objetos inanimados e coisas, possuem uma forma substancial. No caso de coisas vivas – e isto inclui todas as coisas vivas – Aristóteles se refere às suas formas substanciais como sua *psuche*, um termo que é frequentemente entendido como sugestão de mente ou alma. (Em relação a isso, pense no termo *psyche*, a partir do qual nós temos o nome da disciplina psicologia, tradicionalmente concebida como o estudo da mente.)

A proposta – de que coisas vivas, assim como objetos inanimados, possuem uma forma substancial – é infeliz por duas razões. Primeiro, ela pode parecer implicar que tem afinidades com o conceito cartesiano de alma como uma substância incorpórea, capaz de existência independente, enquanto esta muito enfaticamente não deve ser concebida como uma substância lógica, mas sim como uma capacidade ou conjunto de capacidades que logicamente podem existir apenas como as características das coisas das quais são capacidades. Em segundo lugar, enquanto para Descartes apenas os seres humanos possuem almas, todas as coisas vivas, incluindo até mesmo plantas e vegetais, possuem *psuchai* (o plural de *psuche*), de acordo com Aristóteles. Em seu significado original, *psuche* significa sopro, ou sopro vital, portanto retendo seus vínculos com o vento e o poder vital. O sopro vital de alguém é o último suspiro que alguém exala no momento da morte. Aristóteles

não toma o termo *psuche* em seu significado literal, mas o considera, em vez disso, como o princípio da vida animal e vegetal. Portanto, *psuche* não é identificada com a consciência, ou com ela relacionada, como a mente ou a alma são em Descartes, e nem tem qualquer significado religioso. No *De Anima*, Aristóteles escreve:

Agora tem sido dito em geral o que a alma é: é a essência definidora de uma coisa, o “que isto é para ser o que é” de um corpo da espécie descrita [isto é, um corpo natural que tem vida]. Agora, suponha que algum instrumento, tal como um machado, fosse um corpo natural; sua essência seria ser “o que isto é para ser um machado”, e isto seria sua alma. Se isso lhe fosse removido, ele deixaria de ser um machado, exceto no nome.<sup>3</sup>

Outra ilustração de *psuche* é apresentada pela bem conhecida cena do *Monty Python* em que John Cleese vai comprar um papagaio, e dão-lhe um papagaio empalhado e pregado em seu poleiro. Mas um papagaio empalhado não é o que ele queria – ele queria um pássaro vivo, capaz de imitar vozes, beliscar a orelha das outras pessoas, matraquear à porta de sua gaiola, voar em volta da sala, ser capaz, em síntese, de fazer todas as coisas que papagaios caracteristicamente têm o poder de fazer. Um papagaio morto e empalhado não é mais um papagaio, assim como um balão estourado não é mais um balão: ambas as coisas perderam sua forma substancial, no caso do papagaio sua *psuche*, as capacidades que faziam deles os tipos de coisas que eles eram. De maneira semelhante, ainda que nós possamos nos referir a um cadáver como uma pessoa morta, aqui a palavra “pessoa” é utilizada apenas como homônimo – tem o mesmo som de quando aplicada a um ser humano vivo, mas estritamente não tem aplicação para um corpo morto. Uma pes-

soa adormecida pode parecer uma pessoa morta, mas a pessoa dormindo retém a capacidade de acordar, pensar, comer, conduzir discussões com amigos e sentir dor, enquanto um cadáver perdeu permanentemente a capacidade para qualquer dessas coisas.

### *Psuche e a hierarquia de capacidades*

Dado que para Aristóteles a posse da *psuche* é o conjunto de capacidades que uma coisa viva tem que a faz ser o tipo de coisa que ela é, mesmo os invertebrados inferiores, assim como os vertebrados, têm uma *psuche* ou alma, assim como de fato também o têm as plantas. A alma de uma planta é sua capacidade de nutrir-se a si mesma, uma “alma nutritiva”, como formula Aristóteles. Essa pode soar estranha para qualquer um educado na tradição platônica/cartesiana, de acordo com a qual uma alma é uma misteriosa substância lógica incorpórea. As almas dos animais, isto é, suas capacidades, vão muito além daquelas das plantas. As plantas em geral não podem se mover, apesar de existirem exceções, tais como a espécie de mimosa cujas folhas se enrolam quando tocadas, ou a dionéia, que tem uma armadilha em miniatura semelhante a uma mandíbula que fecha quando da entrada de um inseto. Mesmo aqui, no entanto, me parece ser o caso de que esses movimentos das plantas são meramente haptotrópicos – isto é, iniciado pelo toque – e não devem ser explicados em termos de crenças e propósitos. Eles não resultam, em outras palavras, em ações genuínas que possam ser empreendidas com um propósito.

Ao contrário, animais que genuinamente podem agir precisam da percepção para se orientar em seu ambiente, por

exemplo, para perseguir objetivos tais como evitar o perigo e buscar os meios para sua sobrevivência e procriação. Dos animais se diz que eles possuem uma *alma sensitiva* que pressupõe a posse de uma alma nutritiva. Uma vez que aos animais falta a linguagem, eles não podem formar uma concepção dos fins que adotam como fins. Entretanto, desejos simples e crenças podem ser a eles atribuídos. Suas percepções são igualmente restritas. Faltando a linguagem, eles não podem fazer uma avaliação a respeito do modo como as coisas se apresentam e de como realmente são. Eles percebem o que Aristóteles chama de “sensíveis especiais”: a eles faltam crenças gerais ou universais e retêm “apenas uma impressão e memória de particulares”, como o cachorro que pode lembrar onde o osso está enterrado, mas, não possuindo os conceitos que vêm com a linguagem, não é consciente disso como algo que possa ser classificado biologicamente como um osso, ou como uma coisa a ser desenterrada na hora apropriada.

### A *psuche* humana

No topo da hierarquia das coisas vivas dotadas de *psuche* está o ser humano. O que é distintivo dos seres humanos como espécie é sua capacidade para a racionalidade, o que tipicamente inclui o poder de formar julgamentos que podem ser verdadeiros ou falsos; ser capaz de raciocinar logicamente e detectar inconsistências lógicas e *non sequiturs*; ter a capacidade de fornecer razões para crenças e cursos de ação; e fornecer justificativas morais para o que eles estão fazendo, e para suas reações emocionais. A lista de atividades não busca ser exaustiva, mas apenas lembrá-lo de alguns dos aspectos da racionalidade humana. Portanto, dos seres humanos é dito que eles possuem

uma *alma racional*. A capacidade de adquirir e usar a linguagem é crucial para a posse e o desenvolvimento da racionalidade. Como sugeri anteriormente, isso não apenas torna possível a formação de uma atitude proposicional, mas é constitutivo do pensamento racional que requer a manipulação de proposições para que ele aconteça. O próprio Aristóteles não discute explicitamente a importância da linguagem para que haja a possibilidade do pensamento racional. Ele se preocupa com o pensamento e a razão, e não com a linguagem como tal, mas a importância da linguagem para a racionalidade é produto de uma judiciosa interpretação de Aristóteles.

Com a linguagem também vem a capacidade de autoconsciência, e consequentemente uma avaliação racional e ajuste de crenças, valores, ações e políticas de que falei anteriormente no Capítulo 1, seção “O caráter holístico dos estados intencionais e sua relação com o estado de consciência”. Segundo Anthony Kenny,<sup>4</sup> a *psuche* humana é mais bem pensada como uma *capacidade de segunda ordem*: é a capacidade de adquirir uma gama de outras *capacidades de primeira ordem*, as várias atividades compreendidas racionalmente, figurando preeminente entre elas. O aprendizado de línguas nitidamente ilustra no que resulta a alegação de Kenny. Agora eu não tenho a capacidade de falar fluentemente tagalo (uma língua filipina), mas por meio de estudo diligente e prática eu poderia adquirir essa capacidade. Igualmente, eu provavelmente poderia adquirir, da mesma maneira, a capacidade de falar fluentemente francês ou alemão. Qual é o limite da capacidade de segunda ordem para adquirir as capacidades de primeira ordem para falar várias línguas, porém, vai claramente variar de indivíduo a indivíduo. Obviamente, seres humanos normais podem adquirir muitas

outras capacidades de primeira ordem: ser capaz de andar de bicicleta, jogar golfe, tocar piano, escrever livros, cozinhar refeições suntuosas, e assim por diante.

Na medida em que nós descendemos na árvore evolucionária, os tipos de capacidades que as coisas vivas possuem mudam: muitos animais não-humanos possuem capacidades que nós como humanos não possuímos – ser capaz de correr a 90 quilômetros por hora, por exemplo, como um guepardo, ou orientar-se por ecolocalização como um morcego. Plantas, também, podem possuir capacidades de que seres humanos não dispõem: ser capaz de fazer fotossíntese, ou converter amido em açúcar, por exemplo. Por outro lado, o guepardo não possui a capacidade de aprender línguas e consequentemente também não possui o poder de se engajar em todos os tipos de trocas linguísticas características da vida humana; e as plantas não podem responder quando falamos com elas. É a capacidade distintamente humana de construir linguagens e sistemas simbólicos que garante a preeminência dos humanos sobre o mundo animal, devido em parte ao entendimento e subsequente controle dos fenômenos que a linguagem torna possível.

### **EXERCÍCIO 8.2**

Como Aristóteles faz a distinção entre a *psuche* de plantas, animais e pessoas?

### **Poderia haver pessoas não-humanas?**

Aristóteles diz que a forma ou *psuche* de seres humanos é sempre encontrada em carne e osso, e isso sugere que nada cuja matéria não consista nesses materiais possa contar como humano. Do

ponto de vista do que constitui a espécie humana biológica isso é um truismo. Mas suponha que tomássemos “humano” como outra maneira de falar de pessoas, e então perguntamos onde podem existir pessoas distintas dos seres humanos cujos elementos constituintes e organização interna fossem muito diferentes dos nossos, tal como marcianos, cardassianos, ferenguis e vulcanianos, para tomar emprestados alguns exemplos da ficção científica. Parece não haver razão, no que diz respeito à abordagem de Aristóteles sobre a natureza das pessoas, que elimine esta hipótese, pois Aristóteles enfatiza a importância de várias capacidades específicas das pessoas, e qualquer entidade capaz de executar essas funções seria uma pessoa. É verdade que, até agora, as únicas pessoas de que temos conhecimento são seres humanos, mas o conceito de uma pessoa não é o mesmo que o conceito de um ser humano, e então não há razão lógica para que extensões desses conceitos precisem coincidir necessariamente. Nossa existência como pessoas depende de nossas funções e capacidades, não de nossa composição material que compartilhamos com um cadáver, que não é mais uma pessoa, pois lhe faltam as capacidades requeridas.

Seria prontamente compreensível que tudo isso fossem reminiscências do funcionalismo (veja Capítulo 5, seção “Funcionalismo metafísico”), e é assim porque os aristotélicos modernos como Hilary Putnam e Martha Nussbaum sugeriram que Aristóteles estivesse aderindo ao funcionalismo. Porém, em um texto não publicado, David Fischer salienta que na espécie de funcionalismo forte associado com as obras iniciais de Putnam nessa área alguns aspectos e desenvolvimentos, que foram esboçados no Capítulo 5, não podem ser encontrados em Aristóteles:

Aristóteles é explícito ao afirmar que é tão errado reduzir a mente à forma ou à função quanto tentar reduzi-la à matéria. A raiva, por exemplo, não é redutível nem à matéria nem à forma, mas é ambos: forma na matéria, *logoi enhuloi* [forma materializada].<sup>5</sup>

Isso parece estar no caminho certo, pois vimos, no Capítulo 5, que o funcionalismo não pareceu possuir o poder de acomodar as características fenomenológicas subjetivas de nossa experiência, nem efetivamente pareceu ser capaz de fornecer uma avaliação satisfatória da intencionalidade.

### **Capacidades: seus exercícios e veículos**

Há algumas outras distinções conceituais que merecem ser feitas e exploradas, pois são particularmente esclarecedoras para que possamos compreender como surgiram certas concepções equivocadas e bem conhecidas sobre a natureza das pessoas. Primeiro, uma distinção precisa ser traçada entre uma *capacidade* e o *exercício* dessa capacidade. A capacidade não exercida é referida por Aristóteles como a primeira atualização, por exemplo, ser capaz de ver; e o exercício da capacidade, como uma segunda atualização, isto é, realmente ver.

Capacidades, por sua vez, podem ser *ativas* ou *passivas*. A capacidade do papagaio de imitar pessoas é uma capacidade ativa, mas é claro que esta capacidade não está sendo exercida quando o papagaio permanece calado em seu poleiro. Balões, por sua vez, são consideravelmente passivos: eles simplesmente ficam lá até que alguém ou um sopro de vento façam com que se movam. Porém, ser capaz de se mover com um sopro de vento ainda é uma capacidade que o balão tem, muito

embora, na ausência de qualquer movimento do ar, a capacidade do balão para o movimento não esteja sendo atualizada. Por fim, podemos distinguir entre a capacidade de uma coisa e o *veículo* dessa capacidade. O veículo de uma capacidade é aquela característica, ou características, que explica como a coisa em questão tem essa capacidade. No caso do balão, isso irá consistir em uma membrana elástica não-porosa que pode ser inflada e então selada para reter o gás em seu interior. Destrua o veículo da capacidade, digamos, furando o balão, e você, portanto, destruirá a capacidade. No caso do papagaio, o veículo de sua capacidade de imitar as pessoas será encontrado em algo mais complicado, a saber, na organização interna do sistema neurológico do papagaio.

Compare a capacidade do papagaio para imitar a fala com o poder de um ser humano para falar e verdadeiramente entender as palavras que escuta ou lê. Nós já sabemos bastante sobre o veículo da capacidade por causa da obra de Pierre Broca (1824-1880) e Karl Wernicke (1848-1905). Broca foi um neurocirurgião e patologista e exercia também a prática médica. Leborgne, um de seus pacientes, estava sofrendo do que é conhecido como afasia. O paciente era capaz de pronunciar apenas algumas palavras simples e era incapaz de escrever, apesar de saber o que queria dizer. Quando questionado, respondia com gestos expressivos, mas apenas podia proferir a sílaba "tan" repetidamente. Broca descobriu que Leborgne era incapaz de falar no decurso do tratamento que lhe prestava para uma infecção gangrenosa, mas ele morreu após apenas uma semana de tratamento médico. Broca fez uma autópsia e descobriu ~~uma~~ lesão superficial no lobo frontal ~~esquerdo~~. Observações subsequentes levaram Broca a concluir que o dano à parte posterior da terceira circunvolução

frontal esquerda destruiu a capacidade de discurso articulado. Como resultado, essa área do cérebro é atualmente conhecida como área de Broca. Hoje, a teoria de Broca se confirma por escaneamentos cerebrais. Pode-se observar que a parte do cérebro que ele identificou torna-se ativa quando uma pessoa põe-se a falar. Karl Wernicke, médico e psiquiatra, nasceu no que era Prússia, atualmente Polônia. Wernicke descobriu que, se a parte occipital da área da fala é danificada, uma pessoa não pode entender palavras faladas ou escritas, mas pode falar com velocidade e ritmo normais, sendo o conteúdo da fala uma espécie de jargão sem sentido, destituído de coerência e significado.

Naturalmente, há muito que ainda não sabemos sobre como seres humanos normais têm a capacidade para a linguagem, mas em vista dos futuros avanços promovidos na neurologia temos tudo para acreditar que em princípio isso possa ser descoberto. Agora, para um dualista cartesiano, o veículo da capacidade de ser um usuário de linguagem precisa ser a alma incorpórea, e de fato a capacidade de usar a linguagem foi um dos dois testes de Descartes para a presença da alma. Descartes rejeitou a possibilidade de uma explicação mecânica da capacidade distintivamente humana de poder entender e gerar uma variedade infinita de frases nunca antes encontradas. Portanto, não possuindo almas, de acordo com Descartes, os animais necessariamente não possuíam esse veículo, portanto eram incapazes de usar a linguagem. Descartes não quer rejeitar categoricamente uma abordagem mecânica da posse e do uso da linguagem. No entanto, parece que duvidou que uma máquina pudesse ter uma complexidade tal qual a requerida, ao que parece, para fundamentar a capacidade do uso da linguagem. Porém, o recurso a um veículo incorpóreo para explicar a ca-

pacidade para a linguagem pareceria não nos levar a lugar algum, pois não temos a mínima ideia do que seria isto. Logo, a própria natureza do caso nos impediria definitivamente de fazer tal descoberta por meio de investigação empírica.

Como terminei de dizer, uma falha em entender adequadamente os conceitos aristotélicos há pouco delineados pode levar a erros grosseiros. Portanto, o erro central do behaviorismo é que ele equipara uma capacidade ou uma potência com seu exercício. Minha capacidade de falar tagalo é reduzida à minha fala de tagalo ou, ao menos, a uma propensão ou tendência de falar tagalo se certas circunstâncias fossem satisfeitas. A teoria da identidade mente/cérebro, por outro lado, exemplifica uma forma de reducionismo ontológico, equipara a capacidade de segunda ordem da mente ao seu veículo – a organização interna do cérebro e as ocorrências que nele acontecem e tornam possível tal capacidade. Por fim, o transcendentalismo hipostasia ou reifica a mente; ele a eleva ilegitimamente à condição de uma substância lógica ou primária, isto é, erroneamente confere a condição de entidade a capacidades que apenas podem existir como pertencentes a alguém ou a algo, levando desta maneira ao dualismo de sustâncias.

### **EXERCÍCIO 8.3**

Como um entendimento da natureza de uma capacidade pode nos habilitar a evitar os erros do dualismo, do behaviorismo e da teoria da identidade mente/cérebro?

Como a posse de uma capacidade é explicada por referência a seu veículo? Considere primeiramente alguns exemplos não-psicológicos. Um arco de madeira de raio maior que  $r$  tem a capacidade de permitir

que uma esfera de bronze de raio  $r$  passe através dele, não obstante sua capacidade de impedir um cubo de bronze com o lado  $2r$  de passar através dele. Nesse exemplo, fornecido por Hilary Putnam, a explicação não virá da física, mas da geometria, e irá recorrer às "características formais gerais ou estruturais, características que podem ser dadas em termos de leis geométricas sem mencionar a matéria – ainda que claramente todo cubo ou esfera existentes devam ser produzidos com alguma matéria apropriada".<sup>6</sup> Da mesma forma, por que uma chave tem a capacidade de abrir uma fechadura específica? A resposta está no fato de que a forma e todo o modelo da chave são ajustados pela estrutura e operação dos dispositivos internos que compõem a fechadura, de novo uma avaliação geométrica, desta vez em conjunto com uma explicação da operação física da fechadura.

A forma da filosofia da mente de Aristóteles está começando a aparecer. Para que se possa estabelecer uma abordagem da mente, é essencial uma divisão intelectual do trabalho. A tarefa do filósofo é fornecer uma análise dos conceitos mentais e psicológicos, tais como a natureza de sensação e percepção, ação e volição, emoções, habilidades cognitivas e consciência. O trabalho do cientista, do neurologista, é investigar empiricamente as reais estruturas neurais e os processos que tornam possíveis os tipos de habilidades psicológicas delineadas. Nenhuma parte deve tentar trocar, ou ultrapassar, o trabalho da outra. O filósofo, como filósofo, não tem nada a dizer sobre neurologia, visto que a filosofia não é uma ciência empírica, mas comprometida com raciocínios analíticos e *a priori*, investigando relações lógicas entre conceitos. Igualmente, e por oposição, não é tarefa do neurocientista entregar-se à especulação filosófica, porque sem habilidade em filosofia isso é passível de levar a numerosas confusões

conceituais e a equívocos.<sup>7</sup> O programa de pesquisa proposto – que possibilita tanto a abordagem contrastiva quanto a complementar do estudo da mente – pode de fato ser encontrado embrionariamente no próprio Aristóteles em algumas observações que ele faz sobre a raiva:

O filósofo natural [o cientista] e o dialético [o filósofo] dariam uma definição diferente para cada uma das afecções [as emoções], por exemplo em resposta à questão "o que é raiva?". Pois o dialético diria que ela é um desejo de vingança ou algo assim, enquanto o filósofo natural diria que ela é uma ebullição do sangue e uma coisa quente sobre o coração. E, desses, um exporia a matéria, o outro a forma e a fundamentação.<sup>8</sup>

Vale a pena comparar essa passagem com o que Peter Hacker tem a dizer sobre nossa habilidade de dar razões racionais para nossas emoções em oposição ao fornecimento de uma explicação dos processos neurológicos que possibilitam tal dar razões:

Uma emoção pode certamente "resultar" em resposta a uma causa. Inquestionavelmente, o funcionamento apropriado do cérebro é uma condição causal de sentirmos emoções. Dano aos córtices pré-frontais ventromediais, por exemplo, ou aos córtices somatossensórios no hemisfério direito, compromete gravemente a capacidade de resposta emocional normal. Mas seria equivocado equiparar as condições causais para a possibilidade de sentir uma emoção à causa de uma emoção em um dado momento. Então, por exemplo, o que atemoriza alguém (por exemplo, o som de um tiro) é a causa de seu medo, não a condição de seu cérebro que torna o medo possível. Além disso, o objeto de uma emoção não precisa ser o mesmo que (apesar de poder muitas vezes coincidir com) a causa da emoção. Aquilo que atemoriza alguém – a saber, o som de

um tiro – não precisa ser aquilo do que alguém tem medo – a saber, ser morto –, assim como o que torna Otelo ciumento (as mentiras de Iago) não precisa ser aquilo de que ele tem ciúmes (o suposto amor de Desdêmona por Cássio).<sup>9</sup>

O mestre de Aristóteles, Platão, fez uma distinção similar quanto a dar razões para a ação de alguém no *Fédon*. Esse caso é discutido por Godfrey Vesey em um livro escrito na forma de um diálogo com Anthony Flew: *Agency and Necessity*<sup>\*</sup>. Sócrates já foi condenado à morte por ingestão de cicuta e tem a escolha de fugir ou esperar pela chegada daquele que transporta o veneno. Vesey comenta:

*Sócrates senta para esperar o portador do veneno. A questão é: qual é a causa de seu sentar-se?*

*Sócrates, no diálogo, compara duas respostas diferentes:*

- a) A razão pela qual estou sentado aqui agora é que meu corpo é composto de ossos e tendões, e como os ossos se movem livremente em suas juntas, os tendões, relaxando e contraíndo, me habilitam de alguma maneira a dobrar meus membros, e esta é a causa de eu estar sentado aqui em uma posição curvada.
- b) Visto que Atenas pensou ser melhor me condenar, portanto, de minha parte, pensei que é melhor sentar aqui, e mais correto ficar e me submeter a qualquer penalidade que ela ordene. Porque, pelo cão egípcio\*\*, eu fantasia que estes ossos e tendões estariam nas vizinhanças de Megara ou Boécia há muito tempo se eu não pensasse que seria

mais correto e honrado submeter-me ao que quer que seja que meu país ordene do que correr e fugir.

*Sócrates diz que a resposta do segundo tipo (b) fornece a verdadeira razão para se sentar ou deitar:*

Se fosse dito que sem ossos e tendões eu não deveria ser capaz de fazer o que penso que é correto, isso seria verdadeiro. Mas dizer que é por causa deles que faço o que estou fazendo, e não pela escolha do que é melhor, seria uma forma de expressão muito frouxa e inexata. A imaginação não tem condições de distinguir entre a causa de uma coisa e a condição sem a qual ela não poderia ser uma causa.<sup>10</sup>

#### EXERCÍCIO 8.4

Resuma os diferentes, porém complementares, papéis adotados pelos filósofos e neurologistas no estudo da natureza da mente.

#### Explicando a consciência

Como vimos na seção anterior, podemos facilmente explicar como uma chave possui a capacidade de abrir a fechadura à qual se encaixa. Mas podemos explicar como as estruturas e os processos neurológicos contêm o veículo para a posse e o exercício da consciência? O próprio Aristóteles nunca menciona o termo\*\*\*, mas isto não é surpreendente, pois, como Wilkes nos lembra, o verbo “ser consciente” surge por volta 1620, e o substantivo “consciência” não aparece até 1678. Além disso, acrescenta:

\* N. de T.: Em português: Ação e necessidade.

\*\* N. de T.: Referência socrática a Anúbis, Deus egípcio que representa o Mercúrio dos sábios, o Fluido, o Ar e a Água sob uma só forma.

\*\*\* N. de T.: No original: *word*.

No século XX, temos pouca dificuldade em compreender as palavras dos antigos gregos que eram “dramaturgos psicológicos”, como Sófocles, Eurípedes e Menandro, e ainda – como Hamlyn corretamente observa – e eles tampouco tinham preocupação com a consciência; mas o ponto principal é que ninguém acredita que esses escritores [...] tenham “deixado qualquer coisa de fora”. A conclusão é que qualquer coisa dita em termos de consciência pode ser dita igualmente bem sem ela.<sup>11</sup>

Porém, para retornar à nossa questão original, por que a explicação da consciência por referência a processos materiais complexos deveria ser considerada como problemática por alguns pensadores? A resposta quase certamente reside em algo que encontramos antes, a saber, que a consciência parece ser um fenômeno tão inteiramente diferente dos fenômenos físicos espacialmente ordenados. (Compare o comentário de Leibniz na “Introdução” ao Capítulo 1, p. 14, assim como os de McGinn e Chalmers no Capítulo 6, seção “Como o cérebro físico suscita aspectos mentais?”, p. 174-177). Mas, para reiterar o comentário de Hume (veja a mesma seção do Capítulo 6, p. 177), *a priori* qualquer coisa pode ser a causa de qualquer coisa e, como Espinosa salientou, ninguém até agora estabeleceu os limites da potencialidade do corpo; então, em última análise, não há impedimento real algum de invocar a neurofisiologia para explicar a consciência.

O próprio Hacker descarta o problema como não tendo saída, ao escrever:

Por que deveria o fato de seres humanos (e outras criaturas também) desfrutarem de ampla variedade de experiências ser pensado como constitutivo de um mistério? Por que isso deve ser o problema central da neurociência cerebral? É impressionante (e talvez deva ser perturbador)

que a razão mais fundamental seja filosófica, certamente cartesiana. Visto de uma perspectiva particular, pode parecer profundamente enigmático como transações causais no mundo material podem dar origem a qualquer coisa categorialmente distinta da matéria como experiência. Como pode o impacto da radiação sobre cones e bastonetes da retina gerar a experiência de ver algo (conscientemente)? Como podem eventos no córtex “dar origem” a estados mentais conscientes? A perspectiva particular em questão é claramente a cartesiana, de acordo com, a qual o mundo consiste em dois domínios categorialmente distintos, o material (essencialmente caracterizado em termos de extensão) e o mental (essencialmente caracterizado em termos de consciência). Mas essa perspectiva, isto agora deve ser evidente, não é nem necessária (ela não era parte da concepção aristotélica), tampouco não-problemática.<sup>12</sup>

A crítica de Hacker é bem-sucedida, apesar de ser necessário lembrar que sua adoção da visão behaviorista, ou quase-behaviorista, de Wittgenstein quanto à natureza dos estados mentais o auxiliou a desfazer a ilusão da existência de um abismo intransponível entre a consciência e os processos cerebrais (veja Capítulo 9, seção “A abordagem criterial”, p. 236).

### A doutrina do intelecto ativo

Anteriormente, na seção “Capacidades: seus exercícios e veículos”, fomos advertidos contra o transcendentalismo, o erro de reificar capacidades mentais. Infelizmente, e inconsistentemente com tudo mais que ele escreveu no *De Anima*, isso é exatamente o que Aristóteles faz em sua discussão do intelecto ativo. Todas as outras faculdades mentais, incluindo, sobretudo, percepção, emoção e imaginação, precisam de uma corporificação material

para que haja uma unidade sem emendas de forma e matéria. Por outro lado, a intelecção pura, que ele sustenta ser pressuposta por todas as variedades de pensamento, não precisa de qualquer corporificação, e não é de nenhuma maneira dependente de um corpo. No *De Anima* Aristóteles escreve:

Sequer é o caso de que o intelecto esteja agora pensando, agora não. É, além disso, em seu estado separado que o intelecto é exatamente aquilo que ele é; e é nessa solidão que é imortal e eterno, apesar de não termos memória, pois o intelecto separado é não-afetado, enquanto o intelecto que é afetado é perecível, e em qualquer caso não pensa nada sem o outro.<sup>13</sup>

Uma linha similar de pensamento pode ser encontrada no *Generation of Animals*\* de Aristóteles: "Consequentemente perdura que o pensamento sozinho provém do exterior, e que apenas ele é divino; pois a realidade corporal não possui nenhuma conexão com a realidade do pensamento".<sup>14</sup> Ainda no *De Anima* Aristóteles observa:

Se precisamos, então, dizer algo em geral sobre todos os tipos de alma, isto seria a primeira realização de um corpo natural com órgãos. Não deveríamos, então, indagar se a alma ou o corpo constituem uma coisa, assim como a cera e sua marca o são, ou em geral se a matéria de cada coisa é uma com aquilo do que ela é a matéria?<sup>15</sup>

Qual deveria ser nossa reação a uma inconsistência tão impressionante? Vários comentadores estão inclinados a descartar a doutrina da incorporeidade do intelecto ativo como uma aberração que deveria ser colocada de lado. De fato, Katherine Wilkes comenta que não é capaz de entender *De Anima III. 5*, nem mesmo com

o auxílio da literatura secundária, e lamenta que Aristóteles algum dia o tivesse escrito. O que parece tê-lo motivado é o reconhecimento do fato de que os cinco sentidos têm órgãos dos sentidos associados com cada um deles, mas não há um órgão correspondente ao pensamento. Se isso deveria sugerir que o cérebro é o órgão do pensamento, a resposta de Aristóteles seria invocar a natureza pura\*\* do pensamento: a ideia aqui é que, se o intelecto possuísse um traço material que lhe fosse próprio, isso impediria a recepção de qualquer coisa que lhe viesse de fora. Ele escreve:

É necessário, então, uma vez que o intelecto pensa todas as coisas, que seja puro, "que ele possa dirigir", como diz Anaxágoras, que é uma maneira de dizer que ele assim é para que possa ter cognição. Pois uma aparência interna impediria e obstruiria isso a partir do exterior – de tal maneira que ele não pode ter qualquer natureza que lhe seja própria, exceto apenas esta, ser potencial.<sup>16</sup>

Katherine Wilkes utiliza uma analogia esclarecedora para ilustrar o argumento de Aristóteles. Um olho cujo corpo vítreo fosse azul seria incapaz de ver coisas azuis. Para ser capaz de ver todas as cores, o corpo vítreo precisa ser incolor. Similarmente, se o intelecto deve ser capaz de contemplar todas as coisas, ele não pode ter uma natureza física, pois isso iria impedi-lo de considerar a si mesmo. Portanto, precisa ser não-físico e não possuir qualquer corporificação material. Todavia, apesar da ingenuidade do pensamento de Aristóteles nesse momento, o fato da inconsistência com o resto do *De Anima* permanece, então eu também estou inclinado a descartar a doutrina do intelecto ativo como uma aberração.

\* N. de T.: Em português: *A geração dos animais*.

\*\* N. de T.: No original: *unmixed*.

Há outro aspecto do intelecto ativo que precisa de breve menção antes de finalmente encerrarmos este tópico: não está nada claro que de sua alegada capacidade de separação do corpo e de sua natureza eterna resultaria a sobrevivência pessoal do sujeito pensante. Wilkes comenta:

É difícil considerar o intelecto ativo como ele próprio sendo uma forma de pensamento; antes, parece ser o que torna o pensamento possível. Se de fato é isso que Aristóteles diz ser "imortal e eterno", é difícil tornar-se pessoalmente comprometido com a imortalidade, ou por ela animado: não seria nada como eu o que sobreviveria à morte corporal. Talvez a melhor comparação aqui fosse a de Heráclito: a *psuche* "flamejante", após a morte do corpo, une-se à fogueira cósmica e desfruta da eterna rotação descrita em sua cosmologia, mas o pedaço da chama ou éter que foi a *psuche* de Heráclito não é mais Heráclito.<sup>17</sup>

### Pontos fortes da *psuche*

Vimos como a abordagem de Aristóteles torna possível a harmonia entre a análise filosófica e a neurociência, oferecendo um entendimento da natureza da mente. Mas há ao menos três vantagens adicionais que são apontadas a seguir.

#### *Evitando a objeção das muitas mentes*

Um problema posto pela concepção cartesiana da pessoa como alma incorpórea era de que não estava nada claro como as almas deveriam ser individuadas e reidentificadas ao longo do tempo. A abordagem de Aristóteles tem a virtude de evitar essa dificuldade. No caso normal, uma pessoa será dita separada de outra simplesmente pela contagem de

seres humanos corpóreos. Não obstante quão similares na aparência e na conduta duas ou mais pessoas sejam, sua ocupação singular de um dado espaço a um dado tempo fornece um critério inabalável para sua individuação. A *psuche* de tais seres, que comprehende suas capacidades distintas, irá então dever logicamente sua identidade à pessoa cuja *psuche* ela é, assim como o sorriso no rosto do transeunte sem objetivo no Capítulo 2, seção "Dualismo substancial", não sendo uma substância lógica, deve sua identidade àquela do transeunte.

#### *Interação mente/corpo*

O problema de Descartes sobre como dois tipos de coisas aparentemente diferentes, mentes incorpóreas e corpos, possam possivelmente interagir simplesmente não ocorre em Aristóteles, porque a *psuche* não é uma substância lógica, e isso é um avanço significativo da teoria. Como David Fisher argumenta: "Corpo e mente não são duas entidades distintas, mas dois aspectos de uma entidade, tão intimamente relacionados como a cera e o sinete impresso na cera. Nós somos mente incorporada ou mentes-em-corpo".<sup>18</sup>

Comentando tanto sobre o interacionismo cartesiano quanto sobre a teoria da conexão cerebral, em que a alma supostamente interage com o hemisfério esquerdo e não com a glândula pineal – uma hipótese posta em evidência pelo eminentne neurologista Sir John Eccles (1903-1997) em um livro escrito em colaboração com o filósofo Sir Karl Popper (1902-1994)<sup>19</sup> –, Peter Hacker observa:

Ambos os pensadores [Descartes e Eccles] erraram ao conceber a mente como uma entidade de algum tipo. Se tivessem prestado atenção em Aristóteles, pensando a mente como um conjunto de poderes ou

potencialidades, eles teriam estado muito mais próximos da verdade, e não teriam se enredado em problemas insolúveis sobre a interação. Pois evidentemente não faz qualquer sentido perguntar como a habilidade de alguém para fazer as várias coisas que ele pode fazer interagem com seu cérebro.<sup>20</sup>

#### *A falácia mereológica*

A mereologia lida com a lógica das relações entre partes e os todos dos quais elas são partes; a falácia mereológica consiste em tratar uma parte como se ela fosse o todo do qual ela é parte. Essa falácia é endêmica, tanto no dualismo cartesiano quanto em sua contraparte, o materialismo degenerado, a teoria da identidade mente/cérebro, pois ambas as teorias evitam vigorosamente atribuir estados mentais à pessoa viva de carne e osso, preferindo em uma instância predicá-los, no caso do dualismo, à alma incorpórea, e no caso da teoria da identidade mente/cérebro, ao cérebro. Aristóteles era muito atento a este equívoco, escrevendo:

Dizemos da alma que ela é triste e feliz, animada e receosa, e que em ambos os casos fica raivosa, percebe e pensa. E todas essas coisas são tomadas como sendo movimentos. Poderiam então ser os fundamentos para pensar que a alma está em movimento. Mas não precisamos dizer isso. Deixe o ser triste, o ser feliz e o pensamento serem movimentos a seu bel prazer e deixe cada um deles estar em movimento, não obstante se o movimento é produzido pela alma, se, por exemplo, estar raivoso ou com medo consiste nos corações estarem sendo movidos de maneira apropriada, e se o pensamento não passa de um movimento desta parte ou de alguma outra [...] então, dizer que é a alma que fica raivosa será como dizer que é a alma que elabora ou constrói uma

casa. Talvez fosse melhor não dizer que a alma se compadece, aprende ou pensa, mas que o homem o faz em virtude da alma.<sup>21</sup>

Um interessante exemplo da falácia, que aclara o erro central do materialismo redutivo, é fornecido pelo exemplo de um relógio.<sup>22</sup> Um relógio não pode dizer a hora a menos que tenha um mecanismo interno que funcione da maneira correta. Mas não é o mecanismo do relógio que diz a hora, mas os próprios relógios, e podem fazer isto apenas porque têm ponteiros e mostradores que são uma expressão de convenções humanas que têm evoluído e se harmonizado como uma maneira de medir e representar o tempo. Sem a existência de calendários e períodos do dia, resultados de costumes e convenções em um mundo social humano, relógios não seriam possíveis. Consequentemente, é impossível oferecer uma única explicação científica mecânica de como os relógios têm a capacidade de dizer as horas. O erro seria ampliado se fosse sustentado que tudo o que um relógio realmente é são as engrenagens, alavancas e molas dentro dele; ou também, ainda mais radicalmente, insistir que todo o falar sobre relógios e horas deve ser eliminado em prol de descrições físicas das operações dos mecanismos do relógio. Também as tentativas de reduzir um relógio a seu mecanismo, ou eliminar o falar sobre períodos do dia e substituí-lo por descrições de operações internas de relógios são erros precisamente análogos aos da teoria da identidade mente/cérebro e do materialismo eliminatório, respectivamente.

Não obstante, em boa parte da moderna filosofia da mente e em outras partes, a falácia mereológica persiste. Portanto Searle, em *Minds, Brains and Science*<sup>\*</sup>, se pergunta como essa massa cinza-esbran-

\* N. de T.: Em português: *Mentes, cérebros e ciência*.

quiçada dentro de nossos crânios poderia concebivelmente ser consciente, e como processos neurológicos poderiam representar estados de coisas em um mundo externo.<sup>23</sup> De maneira similar, encontramos em um estudo sobre habilidades que o hemisfério esquerdo *pensa com palavras*, que ele *usa discurso e escrita para descrever materiais visuais*; e que o hemisfério direito *pensa com ilustrações e imagens*, que ele *memoriza por meio da ação e das imagens*, que *grava visualmente* e que é o centro para a informação guardada e lembrada. Mas essas especulações e asserções são em grande parte literalmente sem sentido, pois não faz sentido atribuir atividades mentais ao cérebro considerado como um todo, e muito menos a partes do cérebro. São os seres humanos vivos em um mundo público, as conhecidas figuras de carne e osso que encontramos e com as quais conversamos, que falam sobre amigos ausentes, desenham mapas de cidades e países, pintam retratos, teorizam sobre eventos históricos, fazem esculturas, deixam símbolos e marcas e expressam crenças em uma linguagem compartilhada, não cérebros ou seus hemisférios, e isso Aristóteles efetivamente viu. Além disso, como Wittgenstein muito depois veio a salientar, o uso de predicados mentais para atribuir estados mentais a outros assim como a nós mesmos é governado pelos critérios comportamentais que estão em um mundo público e social, não pelo que ocorre dentro dos crânios das pessoas. Nem podem tais predicados adquirir um significado e um uso por referência a estados imateriais de uma alma logicamente privada, pelo motivo de que não poderia haver, em princípio, nenhuma maneira de determinar se tais predicados são aplicados correta ou incorretamente por referência a uma regra.

A doutrina da superveniência também não escapa à crítica de Aristóteles de que estados mentais devem ser predicados da pessoa como um todo e não de alguma parte dela, pois ainda que os estados mentais não sejam mais identificados com estados cerebrais materiais, eles efetivamente dependem destes para sua existência e caráter, é parecer natural interpretar isso como afirmando que eles são, ainda que em um sentido secundário, atributos do cérebro. Certamente Searle endossa essa posição, como várias passagens de seus livros atestam. Portanto, em *Minds, Brains and Science* nos é dito que fenômenos mentais são causados pela microestrutura do cérebro e por ela realizados, e por isso, *a fortiori*, são, portanto, aspectos do cérebro. Embora Searle estivesse atrasado ao abandonar a redução crua de estados mentais a estados cerebrais, é evidente que ele persiste pensando no cérebro como o verdadeiro sujeito de estados mentais. Por exemplo, em uma obra mais recente, *The Mystery of Consciousness*<sup>\*</sup>, ele escreve: "O cérebro é uma máquina, uma máquina biológica, e ele pode pensar".<sup>24</sup> Ao contrário, de acordo com Aristóteles e Wittgenstein, não é o cérebro que pensa, mas as pessoas vivas; mesmo que possa ser verdadeiro que pessoas não poderiam pensar a menos que possuam cérebros funcionando, ainda disso não decorre que um dos órgãos que as compõem realize o pensamento, ainda que ele seja crucialmente vital, em vez da pessoa considerada como um todo.

Se isso ainda não é evidente, por agora deve ser muitíssimo claro que a falácia mereológica igualmente infecta a linguagem do pensamento hipotético de Fodor – isto é, a teoria de que o cérebro tem sua própria maneira de representar estados de coisas externos a ele por meio da lin-

\* N. de T.: Em português: *O mistério da consciência*.

guagem do mentalês – e variações sobre essa hipótese, tais como a de que o cérebro contém mapas ou modelos do mundo externo. Mapas ou modelos, porém, não são objetos naturais como árvores e pedras, mas são produtos de invenção e convenção humanas. Assim como relógios, também mapas e modelos incorporam convenções humanas que têm surgido e se harmonizado em um mundo público, em que convenções sobre escala, método de projeção, como aspectos naturais e criados pelo homem são representados, e assim por diante. Como Peter Hacker apropriadamente nos lembra:

Não há tal coisa como representar um território em um mapa sem empregar um conjunto particular de convenções de representação, envolvendo um método específico de projeção [...] Então, não há mapas que representem sem convenções de representação. Não há convenções de representação sem um uso, pelas criaturas inteligentes, empregadoras de símbolos, da representação. E para usar uma representação corretamente alguém precisa *conhecer* as convenções de representação, entendê-las, ser capaz de explicá-las, reconhecer equívocos e corrigir ou tomar consciência deles quando eles forem apontados. Se certa disposição de linhas é ou não um mapa não é um aspecto *intríseco* das linhas, nem mesmo um aspecto relacional (isto é, a possibilidade de um mapear 1 : 1), mas um aspecto convencional (isto é, o emprego atual, por uma pessoa, de uma convenção de mapeamento).<sup>25</sup>

Pontos paralelos podem ser feitos sobre a linguagem:

As palavras de uma linguagem têm um uso convencionalmente atribuído, e alguém que tenha dominado a linguagem

deve saber o uso, o emprego correto de uma expressão e ser capaz de distinguir um uso correto de um incorreto. Para que um som ou uma marca no papel tenha um uso, para ser usada de acordo com um padrão de correção, deve haver uma prática de usá-la, tipicamente uma prática social. Pois apenas onde há uma prática de empregar um sinal pode também haver uma atividade de equiparar a aplicação de um sinal com um padrão de correção. Uma vez que sinais têm um significado, um uso, apenas na medida em que há uma convenção, um padrão de correção para sua aplicação, então deve haver uma *possibilidade* de corrigir usos equivocados por referência ao padrão de correção para o uso da expressão que é incorporada em uma explicação do significado. O uso da linguagem é essencialmente uma atividade normativa.<sup>26</sup>

À luz desses *insights*, não é de se admirar que teorias representacionistas da mente que tipicamente invocam uma linguagem de hipótese de pensamento, ou a criação de mapas internos, tenham significantemente falhado em apresentar teorias coerentes, pois o engano não consiste em erros nos detalhes que poderíamos esperar corrigir gradualmente por meio de trabalho lento e paciente, mas em um grande erro que deve ser rejeitado completamente desde o princípio. O erro é típico de um reducionista, o erro científico de supor que apenas entidades, métodos, teorias, descrições e explicações empregadas pelas ciências físicas sejam genuinamente respeitáveis. Por exemplo, é significativo que no *Meaning and Mental Representations*<sup>\*</sup>, de Robert Cummins,<sup>27</sup> a máxima de Quine “filosofia da ciência é filosofia o bastante” é citada com aprovação evidente nas páginas preliminares. Mas Quine está precisamente errado sobre esse ponto: a filosofia da ciência é

\* N. de T.: Em português: *Significado e representações mentais*.

mais exatamente *não* suficiente filosoficamente, e a tentativa de reduzir todo tipo de descrição e explicação a um único tipo fisicalista de explicação apenas pode nos levar a uma confusão da qual devemos sempre lutar desesperadamente para escapar.

### **EXERCÍCIO 8.5**

Explique os pontos fortes que a teoria da mente de Aristóteles tem sobre suas rivais.

### **Questões para reflexão**

1. Que problemas do dualismo e da identidade mente/cérebro a filosofia da mente de Aristóteles evita, e de que maneiras?
2. Como Aristóteles distingue entre uma capacidade, seu exercício e seu veículo?
3. O que é a doutrina de Aristóteles do intelecto ativo e como ela pode ser criticada?
4. Qual é a distinção feita por Aristóteles entre matéria e forma, e como ela ajuda a evitar o problema da interação mente/corpo?
5. Explique a falácia mereológica e como a filosofia da mente de Aristóteles logra evitá-la.

### **Sugestões para leituras aprofundadas**

- BARNES, Jonathan. *Aristotle*. Oxford, Oxford University Press, 1982 fornece uma introdução curta e útil a Aristóteles.
- SMITH, P. e JONES, O. R. *The Philosophy of Mind*. Cambridge, Cambridge University Press, 1986 é

uma obra na qual se encontra uma boa discussão da filosofia da mente de Aristóteles.

- STRAWSON, P. F. *Individuals*. London, Methuen, 1959. Para uma abordagem de inspiração aristotélica da natureza das pessoas, especialmente o capítulo 3.
- KENNY, Anthony. *Will, Freedom and Power*. Oxford, Blackwell, 1975 também é útil, especialmente os capítulos 1 e 2.
- NUSSBAUM, Martha e RORTY, Amelie Oksenberg, eds. *Essays on Aristotle's De Anima*. Oxford, Clarendon Press, 1995 é uma rica fonte de discussão sobre a filosofia da mente de Aristóteles por vários comentadores ilustres.
- BENNETT, M. R. e HACKER, P. M. S. *Philosophical Foundations of Neuroscience* ilustra uma discussão detalhada da falácia mereológica que pode ser vista em muitos aspectos como um ensaio em aristotelismo dos dias de hoje que evita o idioma behaviorista de RYLE, Gilbert, *The Concept of Mind*, mas expande e acrescenta a seu projeto.
- HACKER, P. M. S. "Languages, Minds and Brains", in: Blakemore, Colin e Greenfield, Susan, eds. *Mindwaves*. Oxford, Blackwell, 1987.

## NOTAS

1. STRAWSON, P.F. *Individuals*. London, Methuen, 1959, p. 103-04.
2. WITTGENSTEIN, L. *Philosophical Investigations*. Oxford, Blackwell, 1986, p. 281.
3. ARISTOTLE. *The works of Aristotle translated into English*. SMITH, J.A. e ROSS, W.D., eds. Oxford, Clarendon Press, 1910-52, 412b, p.10-15.
4. KENNY, A. *Will, Freedom and Power*. Oxford, Blackwell, 1975, p. 2.
5. FISHER, D. "The Essential Unity of Mind and Body: A defence of Aristotle's Middle Way" (proferida para a South-West London Gastronomico-Philosophical Society em 1º de julho de 2006; não-publicada).
6. PUTNAM, H. e NUSSBAUM, M. Changing Aristotle's Mind, in: PUTNAM H., ed. *James Conant, Words and Life*. Cambridge, MA, Harvard University Press, 1995, p.26.
7. BENNET, M.R. e HACKER, P.M.S. *Philosophical Foundations of Neuroscience*. Oxford, Blackwell Publishing, 2003. Consulte este livro para uma detalhada discussão sobre os erros filosóficos perpetrados por neurologistas no correr dos anos.
8. ARISTOTLE. *De Anima*. [LAWSON-TANCRED H. trad.] London, Penguin Books, 1986, 403a-b.
9. BENNET, M.R. e HACKER, P.M.S. *Philosophical Foundations of Neuroscience*. Oxford, Blackwell Publishing, 2003, p. 208.
10. FLEW, A. e VESEY, G. *Agency and Necessity*. Oxford, Blackwell, 1987.
11. WILKES, K.V. "Psuche versus the Mind", in: NUSSBAUM M.C. e RORTY A.O., *Essays on Aristotle's De Anima*. Oxford, Clarendon Press, 1995, p. 122.
12. BENNET, M.R. e HACKER, P.M.S. *Philosophical Foundations of Neuroscience*. Oxford, Blackwell Publishing, 2003, p. 262.
13. ARISTOTLE. *De Anima III*. 5, 430a.
14. Apud BARNES, J., in: *Aristotle*. Oxford, OUP, 1982, p. 68.
15. ARISTOTLE. *De Anima II*. 1, 412b.
16. ARISTOTLE. *De Anima III*. 4, 429a.
17. WILKES, K.V. "Psuche versus the Mind", p.125.
18. FISHER, D. "The Essential Unity of Mind and Body", p.3
19. ECCLES, J. e POPPER, K. *The Self and its Brain*. Berlin, Springer Verlag, 1977, p. 357.
20. BENNET, M.R. e HACKER, P.M.S. *Philosophical Foundations of Neuroscience*, p. 52.
21. ARISTOTLE. *De Anima I*. 4, 408b.
22. Sou grato a Peter Hacker por este exemplo que me foi sugerido durante uma conversa com ele.
23. SEARLE, J. *Minds, brains, and science. The 1984 Reith Lectures*. London, Penguin, 1989.
24. SEARLE, J. *The Mystery of Consciousness*. London, Granta Publications, 1987, p. 13.
25. HACKER, P.M.S. "Languages, Minds and Brains", in: BLAKEMORE, C. e GREENFIELD, S., eds. *Mindwaves*. Oxford, Blackwell Publishing, 1987, p. 497-8.
26. Ibid., p. 495-6.
27. CUMMINS, R. *Meaning and Mental Representation*. Cambridge, MA, MIT Press, 1995.

# 9

## O PROBLEMA DE OUTRAS MENTES

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender como surge o problema de outras mentes.
- Entender as dificuldades envolvidas na tentativa de solucionar o problema por meio da inferência dedutiva ou indutiva da existência de tais estados a partir do comportamento de uma pessoa.
- Entender o argumento por analogia para outras mentes, e suas fraquezas;
- Entender o ataque de Wittgenstein à concepção cartesiana de como os termos mentais adquirem seus significados.
- Estar apto a avaliar criticamente a abordagem criterial positiva de Wittgenstein sobre como termos mentais adquirem seus significados e a avaliar sua contribuição para nosso entendimento do problema de outras mentes.
- Estar apto a avaliar criticamente a tentativa de P. F. Strawson de resolver o problema de outras mentes.

### Introdução

A maior parte deste livro, até o presente momento, disse respeito à metafísica da mente. A metafísica aborda a questão sobre que tipos de coisas existem, e no que consiste sua existência ou ser. Assim, uma questão central tem sido qual é a natureza da mente. Mais especificamente, nossa investigação disse respeito a qual é o *status ontológico* dos estados mentais, assim como o dos sujeitos ou possuidores

de tais estados. O foco deste capítulo, por outro lado, está na epistemologia da mente. A epistemologia é o ramo da filosofia que lida com questões como o que é conhecer algo, que tipo de coisas podemos conhecer, e como podemos conhecê-las. O interesse particular deste capítulo é determinar se, e como, cada um de nós pode conhecer a existência de mentes distintas de nossas próprias.

Comecemos recordando que um tema recorrente na filosofia da mente

é a privacidade do mental. Lembre, por exemplo, do argumento segundo o qual se os estados mentais de uma pessoa são radicalmente privados, mas seus estados cerebrais são publicamente acessíveis, então, pela lei de Leibniz (ver Capítulo 2, seção “O argumento da dúvida”), estados mentais não podem ser identificados com estados cerebrais ou a eles reduzidos. Este argumento incorpora uma imagem da mente que em grande medida herdamos de Descartes. Emprego a qualificação “em grande medida” porque me parece que muitas pessoas que nunca chegaram sequer a ouvir falar de Descartes, e não estão familiarizadas com sua filosofia da mente, subscrevem algo como sua visão da mente como um reino radicalmente privado, uma paisagem ou teatro interno, acessível apenas ao possuidor da mente em questão.

Se, no entanto, os estados mentais realmente possuíssem uma privacidade radical, ou uma privacidade lógica, como ela veio a ser chamada, que os protege do escrutínio público, então deve para sempre subsistir uma dúvida de que alguém possa alguma vez realmente saber o que está ocorrendo na mente de outra pessoa, ou mesmo, como caso limite, se outras pessoas de fato possuem uma mente. Ao contrário, não há problema paralelo algum no que tange à existência de outros cérebros. No que diz respeito a eventos físicos, tais como os processos cerebrais, eu poderia em princípio observar o que está ocorrendo em seu cérebro, do mesmo modo que você poderia diretamente observar, ou então descobrir, mediante um eletroencefalograma, que mede o output elétrico do cérebro, o que está ocorrendo no meu. Não possui meio especial algum de descobrir o que está acontecendo em meu cérebro que esteja fora do seu alcance, assim como você não possui meios especiais de descobrir o que está ocorrendo

em seu cérebro que esteja fora do meu alcance. Nossa acesso a nossos próprios processos cerebrais, e o acesso de cada um de nós aos de outrem, é simétrico. Nenhum de nós possui qualquer acesso privilegiado ao que está ocorrendo em nossos cérebros que esteja fora do alcance de outrem. No entanto, se a imagem que Descartes oferece está correta, nosso acesso aos conteúdos de nossas próprias mentes, e o acesso de cada um de nós às mentes de outrem, é assimétrico, uma vez que cada pessoa individual possui um acesso privilegiado que não está disponível a outrem. É a doutrina do acesso privilegiado que abre espaço para a dúvida sobre a existência de outras mentes, uma dúvida que está excluída no caso de outros cérebros, que são diretamente observáveis por todos e cada um.<sup>1</sup>

O problema da existência de outras mentes geralmente não nos chama a atenção na vida cotidiana. Atribuímos estados mentais a outros seres humanos sem hesitar; a possibilidade de que o que consideramos ser outras pessoas pudesse tratar-se, na verdade, de zumbis sem mente não apenas jamais passa pela nossa cabeça, mas seria imediatamente descartado como uma possibilidade ridícula pela maioria das pessoas, caso fosse apresentada. Mas por que estamos tão certos de que outras mentes existem, e, crucialmente, sobre que base?

Tal base, pareceria, só pode ser o comportamento que outras pessoas exibem, a saber, o que elas dizem e fazem. Em meu próprio caso não preciso observar meu comportamento em um espelho para estar apto a saber o que estou pensando e sentindo, e tampouco preciso fazer quaisquer inferências partindo de evidências mais básicas para chegar a meu estado mental. Estou apto a conhecer direta e não-inferencialmente meu próprio estado mental. Mas no seu caso parece que sei, se é que de fato

sei, qual é seu estado mental apenas com base no seu comportamento. De que outro modo eu poderia conhecê-lo, se não a partir da observação de seu comportamento, o estado mental que está por trás dele?<sup>2</sup>

O problema, no entanto, é este. Estados mentais, por um lado, e instâncias de comportamento, por outro, não podem, ao que tudo indica, ser identificados uns com os outros, a despeito dos esforços dos behavioristas analíticos para reduzir estados mentais a padrões de comportamento. Esses esforços, como vimos no Capítulo 4, falharam. Estados mentais teimosamente recusam-se a ser reduzidos ao comportamento, o que leva à conclusão de que ambos podem existir completamente separados um do outro. Do mesmo modo que poderia existir comportamento sem nenhum estado mental correspondente, também poderiam existir estados mentais na ausência de qualquer comportamento. Mas, dado que estados mentais são diferentes e separados de instâncias de comportamento, por meio de que processo de inferência seria possível passar de enunciados descrevendo o comportamento de uma pessoa a enunciados descrevendo seus estados mentais, uma vez que os significados das duas classes de enunciados parecem ser completamente diferentes entre si?<sup>3</sup> Falar sobre o mental, de acordo com o modelo cartesiano da mente como um teatro privado, é falar sobre o que é radical ou logicamente privado. Falar sobre comportamento é falar sobre o que pode ser publicamente testemunhado. Além disso, estados mentais são concebidos como as causas subjacentes do comportamento. Tentar identificar estados mentais com o comportamento que eles deveriam explicar seria como tentar equacionar causas com seus efeitos. Mas, como Hilary Putnam indicou, "causas (dores) não são construções lógicas a partir de seus efeitos (comportamento)".<sup>4</sup>

Se estados mentais pudessem ser identificados com comportamento, não haveria necessidade de inferir a existência de tais estados a partir do comportamento. Nenhum padrão de inferência estaria envolvido ou seria requerido. Ao testemunhar o comportamento de uma pessoa, alguém estaria literalmente testemunhando o estado mental daquela pessoa. Mas, uma vez que os estados mentais não podem ser identificados com o comportamento, parece que alguma espécie de inferência de um para o outro deva estar envolvida. Qual é a natureza dessa inferência?

Há apenas dois tipos de inferência a que podemos recorrer: a inferência dedutiva e a inferência indutiva. O traço distintivo da inferência dedutiva é que, em um argumento dedutivo válido, a conclusão do argumento é logicamente implicada por suas premissas. Isso quer dizer que, se as premissas são asseridas, mas a conclusão é negada, seguir-se-á uma contradição formal. Isso se deve ao fato de que a informação contida na conclusão já está contida nas premissas. Os argumentos dedutivos não são ampliativos: nenhuma informação ou significado podem ser extraídos das conclusões desses argumentos que já não estivessem contidas em suas premissas. Mas isso quer dizer que não podemos inferir dedutivamente os estados mentais de outras pessoas a partir de seus comportamentos. Antes de prosseguir a leitura, considere a questão apresentada no Exercício 9.1.

### **EXERCÍCIO 9.1**

Você consegue entender por que a inferência dedutiva está bloqueada neste caso? Pense sobre isso sozinho por alguns minutos, e depois discuta a questão com alguma outra pessoa antes de ler meu comentário a seguir.

A diferença entre os significados de enunciados comportamentais e aqueles atribuindo estados mentais exclui a possibilidade de passar dedutivamente de enunciados descrevendo comportamento àqueles descrevendo estados mentais. Como vimos, uma vez que falar sobre estados mentais não é equivalente a falar sobre comportamento, qualquer tentativa de nos movermos de premissas puramente comportamentais para uma conclusão atribuindo um item não-comportamental, tal como um estado mental, representaria uma tentativa de injetar mais na conclusão do que o que está contido nas premissas.

### **Indução e o argumento por analogia**

Uma vez que o uso da inferência dedutiva para passar do comportamento aos estados mentais tenha sido excluído, resta-nos a indução. O traço distintivo da inferência indutiva é que a conclusão vai além daquilo que é logicamente implicado pelas premissas. Em outras palavras, esse padrão de inferência é ampliativo. Ele tem a forma “Todos os As observados até o momento foram Bs”, portanto “Todos os As que existem se tornarão Bs”. Claramente há mais informação contida na conclusão do que na (única) premissa. É por isso que asserir a premissa, mas negar a conclusão, não gera uma contradição. A conclusão não contém de uma forma diferente a informação contida na premissa, mas vai além dela, contrariamente ao que acontece no caso dos argumentos dedutivos.<sup>5</sup> Um exemplo de inferência indutiva poderia ser o seguinte: tem-se observado que, se as pessoas que estavam adormecidas são despertadas quando seus olhos se movem rapidamente sob as pálpebras, elas relatam que estiveram sonhando. Supo-

nhamos que se tenha descoberto que isso acontece em todos os casos observados. Com base nisso, faz-se a inferência indutiva: “Sempre que os olhos das pessoas se movem rapidamente enquanto elas estão adormecidas, elas estão sonhando”. Muito claramente, a conclusão apresenta uma alegação muito mais geral do que aquela feita na premissa, que alega apenas que, nos casos observados até o momento, há uma correlação entre os movimentos dos olhos durante o sono e o ato de sonhar.

Como pode a inferência indutiva ser aplicada ao problema de outras mentes? Bem, pode-se argumentar que sempre que certo tipo de comportamento é observado constata-se um mesmo tipo de estado mental ocorrendo simultaneamente, o que fornece base indutiva para a alegação de que sempre que aquele tipo de comportamento ocorre ele será acompanhado pelo mesmo tipo de estado mental. Note, contudo, que, assim como no exemplo do sonho e do movimento rápido dos olhos, para que um argumento indutivo sobre a existência de outras mentes prossiga, é necessário que correlações entre a ocorrência de certo tipo de comportamento e a presença de um certo tipo de estado mental tenham sido anteriormente estabelecidas. Mas como essa correlação é estabelecida, para começo de conversa? Considere essa questão enquanto examina o Exercício 9.2.

### **EXERCÍCIO 9.2**

Você consegue perceber a dificuldade de se estabelecer uma correlação entre a presença de certos tipos de estados mentais e certos padrões de comportamento no caso de outras pessoas?

A dificuldade deve ser razoavelmente óbvia. A correlação não pode ser estabelecida porque os estados mentais de outras

pessoas não podem ser observados. Tudo o que alguém pode observar no caso dos outros é o comportamento deles. Se há estados mentais por trás de tal comportamento é algo que não se pode pressupor. É a existência mesma desses estados que está em pauta.

Há, contudo, apenas um caso em que uma correlação entre estado mental e padrão de comportamento pode ser descoberta, a saber, o meu próprio caso. Suponha que eu perceba que sempre que me sinto triste, choro. Eu então conlubo que sempre que vejo você chorar, a mesma correlação deve valer, porque você é muito similar a mim no que toca à sua aparência externa e comportamento, assim como biologicamente. É apenas no meu próprio caso que estou em contato direto com minha tristeza – como um estado mental interno e privado. No seu caso, posso entrar em contato apenas com metade do fenômeno correlacionado: seu comportamento externo e público. Todavia, por analogia com meu próprio caso, parece indutivamente razoável concluir que o fenômeno público e externo do choro é acompanhado por uma tristeza interna e privada, tanto no seu caso quanto no meu. Esse argumento foi enunciado pelo filósofo John Stuart Mill no século XIX como segue:

Conlubo que outros seres humanos possuem sentimentos assim como eu porque, em primeiro lugar, eles possuem corpos, assim como eu, que sei, em meu próprio caso, que é uma condição que antecede os sentimentos; e porque, em segundo lugar, eles exibem os atos e outros sinais externos, que no meu próprio caso sei por experiência que são causados por sentimentos.<sup>6</sup>

Para tornar isso mais claro, a situação é como aquela em que estou caminhando, em uma rua à noite e vejo sombras nas

cortinas das casas que mostram pessoas aparentemente sentadas para o chá. Eu não posso entrar nessas casas para conferir se aquilo é realmente o que está acontecendo, mas, em um dado momento, chego à minha própria casa. Lá, pela silhueta que vejo na cortina, está minha esposa aparentemente sentada para o chá. No momento seguinte, entro em casa e vejo com meus próprios olhos, diretamente, que ela de fato está sentada para o chá. Tendo estabelecido a correlação entre as sombras na cortina e o que está acontecendo dentro de minha própria casa, eu então conlubo que, por analogia com meu próprio caso, o mesmo deve ser verdadeiro para todas as casas pelas quais passei no caminho para minha própria casa.<sup>7</sup> Com esse exemplo em mente, examine o Exercício 9.3.

### **EXERCÍCIO 9.3**

Você pode perceber uma importante diferença entre o problema das outras mentes e o problema do que está acontecendo nas outras casas? Essa diferença cria um problema para a tentativa de estabelecer a existência de outras mentes por analogia com meu próprio caso?

A ilustração utilizando as cortinas não capta de maneira adequada o problema das outras mentes, pela seguinte razão: se, por algum motivo, eu ficasse em dúvida a respeito de minha inferência indutiva sobre outras pessoas estarem sentadas para o chá, quando as sombras características aparecem nas cortinas, sempre seria possível resolver essa dúvida entrando nas casas das outras pessoas e vendo com meus próprios olhos. Mas isso só é possível porque os interiores das casas das outras pessoas não são privados, no sentido de serem fechados para mim, e desse modo eu poderia ter olhado para

dentro deles caso desejasse. Contudo, se estados mentais, ao contrário do que acontece dentro das casas, são logicamente privados, então nunca posso conferir se minhas inferências indutivas foram acuradas, uma vez que nada jamais poderá ser considerado como a obtenção de um acesso direto aos estados mentais dos outros. Isso tem sido considerado uma séria fraqueza desse tipo de argumento. Eu posso estar seguro da correlação entre certo tipo de estado mental e certo tipo de padrão de comportamento apenas no meu próprio caso. Como poderia, então, concluir tão jubilosamente que o que está acontecendo em todos os outros casos é similar ao que acontece no meu, se meu próprio caso é tudo que tenho para analisar? Parece desenfreadamente irresponsável generalizar a partir de um único caso, o meu próprio, para o de outras pessoas. Além disso, como vimos, nunca tenho a oportunidade de descobrir se minha indução está correta, uma vez que jamais posso vir a conhecer diretamente os estados mentais dos outros.

O argumento por analogia também foi acusado de chauvinismo. Se nós sabemos por analogia com a forma e o comportamento humanos que há mentes distintas das nossas, isso não valerá apenas para o caso dos outros seres humanos, mas também no dos animais, uma vez que estes não se assemelham a seres humanos? Mas então seremos levados, por uma rota diferente, à conclusão de Descartes de que animais não possuem mentes. Há também uma preocupação de que podemos ser incapazes de atribuir mentalidade a seres humanos que são de alguma maneira anormais, e em decorrência disso tratá-los inapropriadamente. A mim parece que essa crítica exagera o caso. Animais, dada a própria natureza do caso, não possuem rostos humanos, por exemplo, mas ainda assim eles efetivamente possuem faces.

Seus membros são diferentes dos nossos, mas ainda assim são reconhecíveis como membros. Do mesmo modo, suas estruturas internas podem diferir das nossas, mas os vários órgãos – coração, fígado, pulmões e, talvez o que seja de maior importância, cérebro e sistema nervoso, são todos reconhecidamente similares aos nossos. Se o argumento por analogia garante que possamos atribuir estados mentais a outras pessoas, então ele garante igualmente que possamos atribuir estados mentais a animais. Não, é claro, o mesmo leque e a mesma complexidade de estados mentais como no caso da espécie humana – não atribuímos, por exemplo, arrependimento a um gato, ou indignação moral a um cão ou a uma lagartixa –, mas isso é exatamente o que deveríamos esperar à luz do leque restrito de comportamentos que os animais exibem em comparação aos seres humanos, e especialmente em vista do fato de que falta aos animais a capacidade de linguagem.

### O ataque de Wittgenstein ao argumento por analogia

Wittgenstein atacou as concepções gêmeas da mente e do significado que o cartesianismo venera. Para Descartes, os termos que descrevem os estados mentais de uma pessoa adquirem seus significados porque nomeiam estados internos, logicamente privados. A natureza do estado nomeado determina o significado da expressão que a ele se refere. Uma vez que apenas a pessoa que possui esses estados pode estar familiarizada com eles, apenas essa pessoa pode conhecer o significado dos vários termos usados para falar sobre aqueles estados. Assim, seria inteiramente possível que o que para mim significa a palavra "dor" e o que para você significa a mesma palavra seja completamente di-

ferente. Como a analogia de Wittgenstein do besouro na caixa deixou claro, a natureza do estado interno nomeado por "dor" em meu caso poderia ser inteiramente diferente da natureza do estado interno nomeado por "dor" no seu caso (ver Capítulo 1, seção "As perspectivas de primeira e terceira pessoas e a alegada privacidade do mental"). Talvez nós jamais viéssemos a descobrir isso, porque eu não posso estar familiarizado com seus estados mentais, e você não pode estar familiarizado com os meus. De fato, se todos os itens que se encontram dentro do escopo de minha experiência são privatamente meus, todas as palavras que uso poderiam ter significados inteiramente diferentes daqueles usados por outras pessoas. O reino dos objetos privados que constituem a experiência deles poderia ser totalmente diferente daquele que constitui a *minha* experiência. Todos nós poderíamos, portanto, estar falando nossas próprias linguagens logicamente privadas, mas nenhum de nós estaria apto a descobrir o que outras pessoas significam com suas palavras, uma vez que essas palavras adquirem seus significados puramente porque nomeiam coisas com as quais apenas os outros podem estar familiarizados. Você deve lembrar que o filósofo do século XVIII, John Locke estava ciente dessa consequência do cartesianismo quando conjecturou que a ideia de uma cor colocada na mente de um homem por uma violeta poderia ser qualitativamente idêntica à ideia de uma cor colocada na mente de outro homem por uma calêndula, e vice-versa. Ambos concordariam que violetas são violetas, e que calêndulas são douradas; contudo, o primeiro significaria com "violeta" o que o outro significaria por "dourado", e vice-versa. Em outras palavras, as cores que eles experenciam quando olham para violetas e calêndulas estariam invertidas, uma em relação à outra, de modo que o

significado das palavras "violeta" e "dourado" para cada um dos homens também estaria invertido, uma vez que as palavras adquirem seus significados particulares a partir das experiências privadas de cada indivíduo. No entanto, se isso é realmente o caso, ou se é meramente uma possibilidade, é algo que jamais poderá ser descoberto, uma vez que nenhuma das pessoas pode exibir à outra os exemplares privados de cor que fornecem o significado dos termos para cores (ver Capítulo 5, seção "Qualia invertidos", nota 10).

Este é o ponto de partida para o argumento por analogia. Eu sei o que signifício com as palavras que uso para falar sobre meus próprios estados mentais, uma vez que eu, unicamente, estou diretamente familiarizado com os estados que fornecem significados aos termos de meu vocabulário mental privado. O problema é determinar se posso alguma vez saber que há outros seres parecidos comigo, que similarmente desfrutam estados mentais e podem falar sobre eles do modo como falo. A essência do ataque de Wittgenstein consiste na tese de que o ponto de partida pressuposto pelo argumento por analogia não é um ponto de partida possível. Eu não posso adquirir um vocabulário mental no ilustre isolamento constituído pelos conteúdos logicamente privados de minha mente, e então tentar argumentar por analogia que outros devem possuir o mesmo que possuo. Começando pelo segundo ponto, uma das frentes do ataque de Wittgenstein se dá nas seguintes linhas. Estou tentando estabelecer, na base de uma analogia, a existência de estados mentais distintos dos meus próprios, mas Wittgenstein questiona o que a expressão "estados mentais distintos dos meus próprios" pode significar, no caso de tais estados encontrarem-se além de qualquer experiência que eu possa deles ter. O problema é paralelo ao quebra-cabeça de

determinar como posso saber que há um mundo externo independentemente existente, se tudo com que posso estar familiarizado é meu próprio mundo privado de imagens mentais ou dados dos sentidos. O mais importante não é que eu não possa saber que há um mundo externo, mas antes saber o que a expressão "mundo externo" significa, se ela não corresponde a nada que eu possa jamais vir a experimentar. Para que conceitos possuam significado, deve ser possível reconhecer ou enunciar aquelas condições que, se satisfeitas, permitiriam a aplicação do conceito ao estado de coisas em questão. Mas se não posso jamais vir a me familiarizar com o mundo externo, ou com outras mentes, não há nada a que tais termos pudessem conceitualmente aplicar-se, e isso pareceria implicar que tais termos são vazios de significado. Conceitos sem intuições, isto é, sem situações de experiência possível às quais eles pudessem ser aplicados, são vazios, segundo a famosa afirmação de Kant na *Crítica da Razão Pura*.<sup>8</sup>

Pode-se pensar que essa objeção possa ser facilmente contornada. Para que possa supor inteligivelmente que alguma outra pessoa esteja com dor, não é preciso que eu esteja diretamente familiarizado com aquela dor. Tudo que preciso é supor que quando aquela pessoa sente dor ela senta o que eu sinto quando sinto dor. Aqui a segunda frente do ataque de Wittgenstein é posta em ação. Ele indica que se aprendi o significado da palavra "dor" apenas a partir de meu próprio caso, como o argumento por analogia pressupõe, então dor significará, com efeito, "sensação que eu tenho". O fato de que sou eu, e não qualquer outra pessoa, que está sentindo a dor entrará na própria definição do que é dor. Mas se isso é verdade então será incoerente supor que outras pessoas tenham dores. Isso não é uma coisa muito fácil de se fazer, diz Wittgenstein, com irônico e de-

liberado eufemismo, pois será equivalente a tentar conceber uma sensação que não sinto a partir do modelo de uma sensação que *sinto*. Se uma parte integral do significado de dor é a sensação que *eu* sinto, então a suposição de que poderiam existir dores não sentidas por mim, mas apenas por outra pessoa, será autocontraditória.

O ponto já tinha sido essencialmente estabelecido pelo grande lógico alemão Gottlob Frege (1848-1925) que escreveu:

Se todo pensamento requer um portador, a cujos conteúdos de consciência ele pertence, então ele seria o pensamento desse portador apenas, e não haveria ciência comum a muitos, na qual muitos pudessem trabalhar. Mas eu, talvez, tenha minha própria ciência, a saber, a totalidade dos pensamentos dos quais sou o portador, e outra pessoa talvez tenha a sua [...] Nenhuma contradição entre as duas ciências seria então possível, e na realidade seria supérfluo disputar sobre a verdade, tão supérfluo, de fato, quase tão ridículo quanto seria duas pessoas discutirem sobre se uma nota de cem marcos é genuína, num caso em que cada uma delas estivesse se referindo àquela nota que possui em seu bolso, e compreendesse a palavra "genuína" em seu próprio sentido particular.<sup>9</sup>

### O problema de outras mentes e a impossibilidade de uma linguagem logicamente privada

Os receios de Wittgenstein relativamente ao argumento por analogia, que acabamos de discutir, representam escaramuças preliminares. O ataque principal está reservado à ideia de que os significados das palavras, incluindo os termos que referem a estados mentais, possam ser radical ou logicamente privados. De acordo

com o cartesiano, posso formar uma concepção de meus próprios estados mentais, posso nomeá-los e fazer um solilóquio silencioso sobre eles, mas não posso nunca estar seguro sobre a existência de estados mentais de outros. Ao contrário, Wittgenstein sustenta que se a mente de fato compreendesse um reino inviolavelmente privado, cujos conteúdos fossem acessíveis apenas à pessoa a quem essa mente pertence, então seria impossível para uma pessoa adquirir e usar predicados mentais, isto é, termos e conceitos gerais que são usados para classificar e descrever a economia mental. Em outras palavras, Wittgenstein está alegando que, de uma perspectiva puramente de primeira pessoa, falar sobre a mente, mesmo sobre a própria mente, não é possível. Atribuições de estados mentais a si mesmo – atribuições em primeira pessoa, como são conhecidas – só são possíveis porque um vocabulário mental foi adquirido em associação com atribuições de estados mentais a outras pessoas, as assim chamadas atribuições de terceira pessoa. Em outras palavras, embora possa parecer que adquirimos e aprendemos a usar um vocabulário mental inicialmente a partir de nossa própria perspectiva privada, de primeira pessoa, um vocabulário que posteriormente passamos a perguntar se pode ser aplicado a outros seres, isso é uma ilusão. Mesmo no caso das palavras que usamos para descrever nossos próprios estados mentais, essas palavras devem ter adquirido seu significado e ter sido aprendidas originalmente em um contexto social essencialmente público (ver Capítulo 1, seção “As perspectivas de primeira e terceira pessoas e a busca de uma abordagem unitária da mente”). A própria possibilidade de atribuição de estados mentais em primeira pessoa, se Wittgenstein está certo, pressupõe atribuição a outrem – isto é, a atribuição de estados mentais a terceiras

pessoas. Conceitos mentais não podem ser privados em sua natureza, e aprendidos e aplicados na privacidade lógica da própria mente de alguém. Se isso está correto, então deve ser defeituosa qualquer abordagem ao problema das outras mentes que, como o argumento por analogia, parte de uma concepção do mental como algo radicalmente privado ao indivíduo, e, desta perspectiva, tente justificar a crença na existência de outras mentes. Antes de revelar os detalhes finos do ataque de Wittgenstein à noção de uma linguagem logicamente privada, pode ser útil fornecer um breve resumo, ponto a ponto, da estratégia geral para atacar a concepção cartesiana do significado discutido anteriormente:

1. O cartesianismo pressupõe que o solipsismo – a visão segundo a qual apenas eu e meus estados mentais existimos – é um estado de coisas possível. Pode não haver um mundo fora da minha mente. Pode ser que existamos apenas eu e os itens mentais privados que constituem minha experiência.
2. Porém, argumenta o cartesiano, mesmo se o solipsismo for verdadeiro, a pessoa pode continuar usando a linguagem para pensar e formular a dúvida cética, a saber, “Não é possível que existamos apenas eu e meus estados mentais?”
3. Porém, isso pressupõe que alguém possa falar uma linguagem que apenas ele possa entender, porque as palavras dessa linguagem se referem aos estados logicamente privados de sua mente.
4. Wittgenstein ataca essa pressuposição. Uma linguagem concebida como a de (3) é uma impossibilidade lógica. As palavras, incluindo os termos para os próprios

estados mentais de alguém, só podem adquirir seus significados por meio de uma referência a coisas em um mundo público, compartilhado.

5. Assim, a própria possibilidade de se formular a dúvida cética, solipsista, implica que deve haver um mundo público para que as palavras que a pessoa usa para formular a dúvida tenham adquirido seus significados em primeiro lugar. Isso significa que termos para estados mentais devem, igualmente, possuir um significado e uma aplicação públicos, no domínio público, implicando assim conhecimento de outras mentes.
6. Consequentemente, o cartesianismo, e a teoria do significado internista e privado, que o primeiro implica, não é um ponto de vista possível de adotar. Mesmo a formulação da perspectiva cartesiana pressupõe sua falsidade.

### **Os detalhes do ataque de Wittgenstein**

Wittgenstein arma duas ofensivas principais contra a ideia de que as palavras, incluindo os nomes de estados mentais, poderiam adquirir seus significados do modo pressuposto pelo cartesianismo. Em primeiro lugar, ataca a noção de uma definição ostensiva e interna que supostamente atribuiria às palavras seus significados. Então, concentra-se na demonstração de que a linguagem, sendo uma atividade guiada por regras, não seria possível caso os significados fossem considerados logicamente privados. Começando com a primeira objeção, observemos mais de perto

o ataque de Wittgenstein à noção de uma definição ostensiva privada.

#### **Definição ostensiva interna**

Definir uma palavra ostensivamente é defini-la por meio de um ato de indicação ou um gesto de apontar. Por exemplo, alguém me pergunta o que significa "grifo"<sup>\*</sup> e eu explico o termo apontando para a imagem de um grifo, como é representado, digamos, no logotipo do Esher College. Ao apontar para o tipo de objeto ao qual se aplica a palavra "grifo", a esperança é de que a outra pessoa venha a entender o que a palavra significa. Palavras não podem ser sempre definidas em termos de outras palavras, pois de outro modo acabaríamos girando em círculo. Deve haver algumas palavras que adquirem seus significados por referência direta ao mundo, a uma realidade não-lingüística, e assim parece que não podemos escapar da necessidade de, em última instância, fornecer definições ostensivas.

Fornecer uma definição do que a palavra "grifo" significa apontando para o logotipo do Esher College é uma atividade que acontece em um mundo público, uma vez que todos podem ver o logotipo. Mas agora suponha que estamos tentando definir um estado mental, como o de uma dor. Se dores, assim como outros estados mentais, realmente são internas e privadas, como o cartesianismo sustenta, então, supostamente, o ato de definição ostensiva supostamente ocorreria no mundo logicamente privado, auto-contido, solitário da minha mente. Eu, por assim dizer, aponto para a dor com o dedo mental de minha concentração, e digo a mim mesmo, com o objetivo de fixar o significado da palavra, "Isto é o

\* N. de T: Animal mitológico com cabeça de águia e garras de leão.

que irei significar no futuro quando usar a palavra 'dor' para descrever para mim mesmo como estou me sentindo". Uma consequência dessa teoria sobre o modo como aprendo a descrever meus próprios estados mentais é que ninguém mais pode descobrir o que eu signifício com "dor". Desse modo, na medida em que forneço a mim mesmo definições ostensivas internas privadas, de meus estados mentais privados, invento minha própria linguagem logicamente privada, uma linguagem que posso entender, mas que outras pessoas não podem.

### **"Tove"\*\* e a impossibilidade de definição ostensiva privada**

Wittgenstein ataca a ideia de que uma definição ostensiva possa ser obtida da maneira radicalmente privada recém-esboçada. Em *The Blue and Brown Books*\*\*, ele nos pede para imaginarmos o que acontece quando se supõe que alguém deva apresentar uma definição ostensiva da palavra "tove" (escolhida porque não é uma palavra em português nem em inglês) apontando para um lápis. O propósito de Wittgenstein é mostrar que o ato de apontar pode ser interpretado de vários modos diferentes; ele poderia significar:

"Isto é um lápis."  
 "Isto é cilíndrico."  
 "Isto é madeira."  
 "Isto é um."  
 "Isto é duro."

Obviamente, alguém poderia perguntar "O que você quer dizer?". Você quer dizer que tove é o formato, ou o

tipo de coisa que isto é, ou apenas que um é tove? Mas para fazer essa questão, e para entender a resposta, a pessoa já deve possuir uma linguagem. De fato, até mesmo para que alguém entenda que essas palavras constituem frases demonstrativas em vez de questões, ou exclamações, ou chamadas por meio de nomes, deve ser feita a suposição de que a pessoa possui uma linguagem. Mas Wittgenstein não deseja fazer essa suposição. Ele está interessado em entender como a linguagem é possível em primeiro lugar. Como uma pessoa adquire o conceito da cor de um objeto em vez de sua figura, ou, alternativamente, o tipo de coisa que é em oposição ao fato de que seja aquele indivíduo particular. E neste ponto é tentador dizer que quando o lingüista privado fornece a si mesmo uma definição ostensiva privada de, digamos, a cor de uma coisa ele apenas se concentra, de alguma maneira, na cor, ou "significa" a cor em oposição ao significado de alguma outra característica. Wittgenstein afirma que somos tentados a pensar na mente como um mecanismo misterioso, não-físico, que pode ocasionar efeitos que nenhum mero mecanismo material poderia suscitar, e que a definição ostensiva é, de acordo com essa imagem, fixada por meio de algum ato mental peculiar, oculto. Wittgenstein escreve:

Aponte para uma folha de papel. – E agora aponte para seu formato – agora para sua cor – agora para seu número (isso soa esquisito). – Como você fez isso? – Você dirá que "significou" uma coisa diferente a cada vez que apontou. E se eu perguntar como isso foi feito, você dirá que concentrou sua atenção na cor, na forma, etc.

\* N. de T.: "Tove" é uma palavra originalmente empregada em um poema absurdo de Lewis Carroll, intitulado "Jabberwocky" (publicado no livro *Alice através do espelho*, de 1871).

\*\* N. de T.: Em português: *Livro azul* e *livro castanho*.

Mas eu pergunto novamente: como isso foi feito?<sup>11</sup>

Antes de prosseguir a leitura, examine as questões apresentadas no Exercício 9.4.

#### **EXERCÍCIO 9.4**

Como a definição ostensiva é possível? Como alguém pode significar a cor em vez do formato? Como os significados de termos para cor e formato podem ser estabelecidos antes de mais nada?

A resposta é, basicamente: por causa das atividades compartilhadas em um mundo social e público. Wittgenstein fornece o seguinte exemplo de como o significado de uma palavra como “azul” poderia ser estabelecido:

Suponha, todavia, que alguém apresentasse a objeção: “Não é verdade que você já deva dominar uma linguagem para que possa entender uma definição ostensiva; tudo que você precisa – obviamente – é saber ou supor para o que a pessoa que está dando a explicação está apontando. Isto é, se ela está apontando, por exemplo, para o formato do objeto, ou para sua cor, ou para seu número, e assim por diante”. – E em que consiste “apontar para o formato” ou “apontar para a cor”? Aponte para uma folha de papel. – E agora aponte para seu formato – agora para sua cor – agora para seu número (isso soa esquisito). – Como você fez isso? – Você dirá que “significou” uma coisa diferente a cada vez que apontou. E se eu perguntar como isso foi feito, você dirá que concentrou sua atenção na cor, na forma, etc. Mas eu pergunto novamente: como isso foi feito?

Suponha que alguém aponte para um vaso e diga “Olhe para aquele maravilhoso azul – o formato não é relevante”. – Ou: “Olhe para o maravilhoso formato – a cor não importa”. Sem dúvida você

fará algo diferente ao responder a cada um desses pedidos. Mas você sempre faz a mesma coisa quando você dirige sua atenção à cor? Imagine vários casos diferentes. Para mencionar uns poucos:

“Este azul é o mesmo que o azul ali adiante? Você percebe alguma diferença?” – Você está misturando tintas e diz “É difícil obter o azul deste céu”.

“Está ficando bom, já se pode ver azul celeste novamente.”

“Veja como esses dois azuis têm efeitos diferentes.”

“Você vê o livro azul que lá está? Traga-o aqui.”

“Este sinal de luz azul significa...”

“Como se chama esse azul? Será ‘anil’?” Você às vezes atenta para a cor erguendo sua mão de modo a ocultar o contorno; ou deixando de olhar para o contorno da coisa; às vezes, fitando o objeto e tentando recordar onde você viu aquela cor antes.

Você atenta para o formato às vezes traçando-o, às vezes contorcendo os olhos de modo a não conseguir ver a cor claramente, e de muitas outras maneiras. Eu gostaria de dizer: Este é o tipo de coisa que acontece enquanto alguém “dirige sua atenção a isto ou aquilo”. Mas não são essas coisas, em si mesmas, que nos fazem dizer que alguém está atentando para o formato, para a cor, e assim por diante. Da mesma maneira que um movimento no xadrez não consiste simplesmente em mover uma peça de tal e tal modo no tabuleiro – e nem, tampouco, nos pensamentos e sentimentos que alguém tem ao efetuar o movimento, mas nas circunstâncias do que chamamos “jogar uma partida de xadrez”, “resolver um problema de xadrez”, e assim por diante.<sup>12</sup>

Um dos pontos que Wittgenstein está apresentando é que essas atividades acontecem em um cenário público. É por meio desses tipos de atividades, e de várias outras maneiras, que a gramática de palavras, tais como termos para cores,

são estabelecidas. Há também uma indicação nessa passagem, que é desenvolvida em outro contexto dos escritos de Wittgenstein, segundo a qual a aquisição de conceitos e linguagem pressupõem um pano de fundo de costumes e atividades. Para começar a apreciar este ponto, considere a situação de alguém que não sabe o que é um rei no xadrez. Pode-se pensar que, para esclarecer isso, tudo que precisamos fazer é pegar um rei e mostrá-lo à pessoa, enquanto apontamos para ele e proferimos as palavras "Isto é um rei no xadrez". Suponha, além disso, que essa pessoa nunca ouviu falar de um jogo anteriormente, menos ainda tê-lo jogado – não apenas xadrez, mas absolutamente qualquer outro jogo. Acho que será óbvio que a tentativa de dizer a ele o que é um rei no xadrez não significará absolutamente nada. Mesmo supondo que ele entenda que estamos tentando dizer a ele como esse objeto é chamado, o máximo que poderia concluir é que um rei no xadrez é uma peça de madeira com determinado formato. Por outro lado, se ele aprendeu jogos simples de tabuleiro por meio de observação, ou talvez experimentando o jogo por conta própria, então quando pegamos o rei e mostramos a ele os vários movimentos que essa peça pode fazer, ele pode gradualmente vir a entender seu uso. Mas isso é apenas porque, como diz Wittgenstein, o terreno para ele adquirir essa compreensão já está preparado. Esta preparação de terreno para o aprendizado do uso de uma peça em um jogo é o que Wittgenstein chama de "preparação do cenário". Ela pressupõe que haja práticas e costumes consistindo em "jogar jogos", e é por meio do contato e envolvimento com esses tipos de atividades que o conceito de diferentes espécies de jogos é adquirido. Baseado nisso, uma compreensão de novos jogos, movimentos e estratégias não familiares pode ser

edificada. Mas nenhum desses recursos ou envolvimentos está disponível para o cartesiano, autoemprisionado no círculo privado de sua própria consciência.

As críticas de Wittgenstein ao modo como o cartesiano supõe que as palavras adquiram significado vão muito fundo, e aqui posso apenas indicar brevemente a direção para a qual elas se encaminham. Deve-se recordar que, para Descartes, e para filósofos como Locke e Hume, que foram fortemente influenciados por ele, a experiência engloba uma série de imagens mentais, asseguradas dentro da fortaleza do reino privado da mente. Esta é uma concepção muito tênue de experiência. Imagens privadas são apresentadas à pessoa, que é essencialmente uma espectadora inerte, passivamente observando o espetáculo momentâneo. As imagens que vêm dos sentidos são fortes e vívidas. A partir dessas imagens, supõe-se que se derivem cópias desmaiadas, e estas constituem conceitos que servem como significados de palavras. Mas essa concepção do significado é profundamente distorcida, e numerosas críticas contra ela foram apresentadas por Wittgenstein. Por razões de espaço, contudo, devo limitar-me a apresentar apenas quatro.

Em primeiro lugar, a mera familiaridade com uma coisa que se enquadra em um conceito não equivale e nem garante a posse do conceito. A familiaridade com um rei particular de xadrez não garantirá por si mesma a aquisição do conceito de um rei de xadrez, uma compreensão do significado da expressão "rei de xadrez". Para adquirir o conceito, como tentei explicar anteriormente, alguém deve ser um agente ativamente envolvido em um mundo público e em suas práticas. O mero registro passivo de imagens não equivale a posse e domínio dos conceitos em questão. Em relação a isso Wittgenstein observa: "Não acredite que

você possui o conceito de cor dentro de si só porque olhou para um objeto colorido – não importando o quanto você olhou. (Não mais do que você possui o conceito de um número negativo por possuir débitos.)".<sup>13</sup>

Em segundo lugar, Wittgenstein argumenta que o significado de uma palavra não pode consistir em um objeto ou coisa, mesmo se pensarmos naquela coisa como um item mental, como uma imagem. Para nos auxiliar a compreender essa crítica, Wittgenstein nos pede para imaginarmos que a imagem mental que supostamente constitui o significado de uma palavra – o conceito que ela expressa – seja substituída por um sinal gráfico. O sinal original – a palavra – está morta, e a imagem supostamente lhe traz vida, soprando nela um significado. Mas um sinal gráfico está tão morto quanto o sinal original. Se em vez de pensarmos no sinal como um desenho pensarmos nele como uma imagem mental, isso não explica miraculosamente como ele pode representar qualquer outra coisa. É apenas do caráter oculto da imagem mental que você precisa, diz Wittgenstein. Ao pensar no conceito do que constitui o significado de uma palavra como sendo uma imagem, você está pensando no significado como se ele fosse apenas um outro objeto que acompanha a palavra. Mas, se o objeto original está morto, o mesmo se aplica ao objeto que você adiciona ao primeiro. O significado de uma palavra não pode, portanto, consistir em um objeto associado, mesmo que seja um objeto mental privado. Precisamos deixar imediatamente de pensar no significado nesses termos e passar a pensar nele, em vez disso, como o uso que dele é feito, a função que ele possui na linguagem, em um mundo público compartilhado.

Uma terceira crítica, relacionada a essa última, mostra que uma imagem

não traz uma interpretação estampada em sua face. Imagens não são auto-interpretativas. Quando olhamos para a imagem na Figura 9.1 (uma imagem do próprio Wittgenstein), o que vemos? O que ela representa? Muitas pessoas dirão que ela representa um homem com um bastão subindo uma colina. Mas será que é isso? Suponha que eu sugira que se trata na verdade de um homem com um bastão deslizando colina abaixo? De fato, pode ser que essa imagem nem mesmo seja de um homem, mas um corte transversal de um pedaço de madeira ou de algum outro material, o cérebro, talvez. Esta última sugestão pode parecer estranha, mas há uma parte do cérebro que se parece com um cavalo-marinho, daí derivando seu nome grego, hipocampo (sua função tem a ver com o armazenamento de memórias de longa duração). Se supusermos que a imagem constitui o significado de uma palavra, como asseriram Descartes e Locke, então não poderemos usar a linguagem para explicar o que a imagem significa, porque isso será argumentar circularmente. O ponto fundamental que Wittgenstein está es-

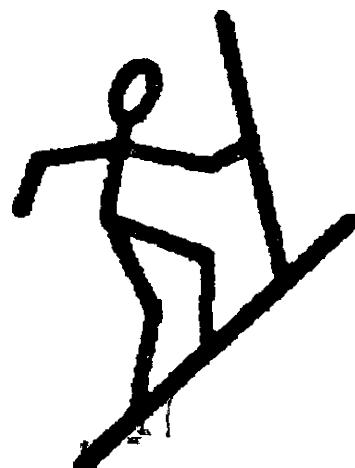


Figura 9.1

tabeleciondo com esse argumento é que supomos que o significado de uma palavra devia fixar ou determinar seu uso, mas uma imagem não pode fixar ou determinar o uso porque sua aplicação não pode ser extraída diretamente dela. Há uma variedade indefinida de maneiras nas quais uma imagem poderia ser usada ou aplicada, e, por conseguinte, uma imagem não pode determinar sua aplicação. No *Blue Book*, Wittgenstein escreve: "Como parte do sistema da linguagem, pode-se dizer, a frase possui vida. Mas somos tentados a imaginar o que dá vida à frase como sendo algo em uma esfera oculta, acompanhando a frase. Mas o que quer que a acompanhe seria apenas mais um sinal".<sup>14</sup>

A imagem com a qual uma palavra está associada, a ideia que é substituída pela palavra na mente, não pode ditar o modo como a palavra deve ser usada, porque a imagem, assim como a palavra com a qual está associada, é ela mesma apenas mais um sinal, e a aplicação de sinais não é intrínseca a eles. Apelar para mais um sinal com o objetivo de esclarecer o modo como a imagem deve ser aplicada meramente recapitula o erro de pensar em sinais como auto-interpretativos. Qualquer interpretação de um sinal, diz Wittgenstein, "permanece solta no ar juntamente com o que ela interpreta, e não pode fornecer-lhe suporte algum. Interpretações, por si mesmas, não determinam o significado".<sup>15</sup>

Como se determina o significado é um assunto volumoso e complexo, no qual não posso entrar aqui. Será suficiente para nossos presentes propósitos reconhecer a conclusão negativa, a saber, que o que quer que o significado seja ele não

pode ser constituído de sinais ou imagens, mesmo se supusermos que esses sinais e imagens existam nos recessos fantasmagóricos e ocultos da mente.

Uma quarta crítica apresentada por Wittgenstein consiste no argumento de que imagens não possuem sintaxe. As frases possuem uma sintaxe que especifica como as palavras podem e como não podem ser combinadas, e em que ordem. Em outras palavras, há regras governando as possibilidades combinatórias das palavras, e o modo como as palavras são combinadas determina o significado da frase em questão. Além disso, a ordem das palavras é importante para determinar se o que foi feito foi uma afirmação ou uma pergunta, como em "Isto é um gato" e "É isto um gato?"\*. Mas que regras poderíamos conceber que pudessem governar as possibilidades combinatórias de imagens, e como poderia a ordem na qual as imagens são organizadas determinar o que elas significam? Uma fileira de imagens poderia significar qualquer coisa, ou coisa alguma. Neste exato momento encontram-se sobre minha mesa os seguintes objetos, organizados nesta ordem: minha luminária, meu escâner, meus óculos, meu monitor, meu relógio, meu mouse e minha impressora. Mas este arranjo não assere ou nega coisa alguma. A experiência cotidiana mostra que se trata apenas de uma organização rotineira de itens que uma pessoa envolvida com edição de textos poderia ter sobre sua mesa. Se eu colocar meus óculos à direita ao invés de à esquerda do monitor, ou se eu os tirar da mesa, que diferença isso fará? Embaralhar os itens na mesa e talvez retirar alguns é precisamente o que não se compara com o ato de fazer mudanças permitidas na ordem das palavras

\* N. de T.: No caso da língua portuguesa, ao contrário do que ocorre no inglês, a diferença entre a afirmação e a interrogação deste exemplo teria se mantido independentemente da ordem das palavras empregadas em cada frase; seria uma questão de entonação.

em uma frase, ou adicioná-las ou subtraí-las, com o objetivo de alterar o significado da frase.

### A impossibilidade de seguir uma regra privada

O propósito da definição ostensiva era fixar o significado de uma palavra. Uma vez que a definição da palavra em questão tenha sido fixada por referência a um exemplar mental privado, uma instância representativa de um tipo particular de coisa, a palavra poderá ser usada repetidamente daí por diante para falar sobre o tipo de coisa em questão. Desse modo as pessoas poderiam confeccionar uma espécie de dicionário na privacidade de suas imaginações. Quando sinto dor, digamos, olho para dentro e, ao apontar o dedo da minha atenção, chamo a sensação de "S". Posteriormente, sinto uma coceira que, de maneira similar, chamo de "T", e ainda mais tarde cócegas que chamo de "U". O Quadro 9.1 apresenta de maneira simples o "dicionário" que é confeccionado dessa forma. A ideia é que quando sinto, digamos,  $\square$  novamente, e pergunto a mim mesmo como devo chamá-lo, posso conferir esse Quadro em minha imaginação, e dizer "Ah, sim – meu estado mental possui o atributo  $\square$ , e isso está correlacionado no dicionário com "S", portanto devo estar presentemente sentindo "S".

Mas, diz Wittgenstein, para que esse procedimento funcione, devo recordar

corretamente o que "S" representa, pois de outro modo usarei "S" equivocadamente. Um sinal que eu pudesse usar da maneira que bem entendesse não pode ser considerado uma palavra. Deve ser possível traçar uma distinção entre o uso correto e incorreto do termo, entre genuinamente seguir uma regra para seu uso e pensar erroneamente que uma regra está sendo seguida. Mas a dificuldade, indica Wittgenstein, é que essas distinções não podem ser traçadas no mundo privado da mente. O que quer que venha a me parecer correto será correto, e isso apenas mostra que não podemos falar sobre algo ser correto, neste caso. Quando uso "S", devo estar apto a recordar corretamente que "S" de fato significa  $\square$ , senão, ao nomear minha sensação presente como "S", estarei enganado. Mas como poderia ser traçada a distinção entre a situação em que recordo corretamente que "S" significa  $\square$ , e aquela na qual meramente me parece que recordo isso corretamente? Se Wittgenstein está certo, essa distinção não pode ser mantida na situação com a qual estamos lidando. No reino privado da imaginação não é possível uma verificação independente de que estou usando a palavra corretamente, de acordo com uma regra. A distinção entre realmente seguir uma regra e apenas parecer seguirla colapsa, e, por conseguinte, não pode ter nenhuma aplicação. Não adiantaria apelar de uma memória para outra para justificar se o uso que alguém fez de um termo foi correto ou não. Isso, observa

**Quadro 9.1** O "dicionário" no mundo privado da mente

| Palavra | Estado mental (item não-linguístico) |
|---------|--------------------------------------|
| S       | $\square$                            |
| T       | ◆                                    |
| U       | ●                                    |

Wittgenstein, seria como comprar várias cópias do mesmo jornal matinal com o objetivo de verificar se a estória que está nele é correta. Sem a possibilidade de uma verificação independente do uso de um termo, tentar definir uma palavra por meio de uma definição ostensiva privada será uma cerimônia vazia, que não consuma coisa alguma. Portanto, a noção de uma linguagem logicamente privada acaba se mostrando, como Wittgenstein diz em outro contexto, um exemplo de contrassenso disfarçado. Isso, por sua vez, significa que o programa cartesiano de dúvida sistemática não pode sequer começar. Descartes não pode partir de dentro do círculo fechado de sua própria consciência privada e daí começar a se perguntar se há um mundo público, independentemente existente. Pois mesmo para que a questão pudesse ser formulada, seria necessário pressupor a existência de um mundo social público, no âmbito do qual apenas a linguagem é possível.

### Como os termos mentais adquirem seus significados?

Como isso tudo se relaciona com o problema de outras mentes? Espero já ter esclarecido parcialmente que o primeiro ponto que deve ser enfatizado é que, de acordo com Wittgenstein, uma pessoa não pode aprender nomes de sensações e outros estados mentais a partir de seu próprio caso. A aquisição da linguagem para falar sobre nossos próprios estados mentais, e mesmo os estados mentais das mentes dos outros, não é possível de uma perspectiva privada, puramente de primeira pessoa. Mas então como ela é possível?

A resposta, não surpreendentemente, é: a partir de um ponto de vista público, de terceira pessoa. Até mesmo a atribuição de estados mentais a si mesmo, feita pelo pró-

prio sujeito, que parece ser fundamentada em uma base puramente privada, de primeira pessoa, só se torna possível porque termos mentais adquirem uma aplicação e um uso essencialmente em uma estrutura de terceira pessoa. Neste estágio isso soa suspeito, pois é muito parecido com o behaviorismo analítico, que sustentava que tudo que dizemos sobre a mente poderia ser traduzido sem remanescentes em termos de demonstrações, reais e possíveis, de comportamento público externo. Mas Wittgenstein nega que ele seja um behaviorista disfarçado. Quais são, então, os detalhes de sua explicação? Ela é constituída de duas linhas principais.

A primeira é a alegação de que o comportamento serve como critério em relação ao qual as atribuições de estados mentais a alguém, tanto a outros quanto a si mesmo, devem ser avaliadas no que toca à sua correção. A outra alegação é que termos mentais não *descrevem* comportamento, mas gradualmente passam a *substituí-lo*. A emissão de gemidos por parte de alguém não é equivalente à produção de uma descrição de sua própria dor; trata-se antes de uma manifestação ou expressão natural daquela dor. Gradualmente, contudo, esta manifestação de dor é substituída, primeiro por exclamações como “ai”, e posteriormente por frases como “Isso dói” ou “Eu estou com dor”. Observemos mais a fundo ambas as alegações antes de tentar fornecer alguma avaliação delas.

### A abordagem criterial

Para começar, a noção de um critério requer alguma explicação. “Um ‘processo interno’, nos diz Wittgenstein, ‘necessita de critérios *externos*’.<sup>16</sup> Muitos comentadores interpretaram isso como significando que termos para estados mentais tornam-se possíveis, isto é, significativos,

apenas porque existem critérios comportamentais governando seu uso. O comportamento fornece a medida contra a qual o uso que uma pessoa faz de palavras mentais pode ser avaliado no que toca à sua correção ou incorreção. Para ser considerado um uso correto, um termo tal como "dor" deve ajustar-se ao resto do comportamento de uma pessoa de maneira adequada. Se me canto, gemo e me contorço, e então anuncio que tenho cócegas, outros podem me corrigir, e indicar que eu deveria ter descrito a mim mesmo como estando com dor em vez de cócegas.

Por contraste com a atribuição de dor a outros, frequentemente se diz que a atribuição de dor a si mesmo é "acriterial". O que isso quer dizer é que atribuo dores a você com base em seu comportamento, mas no meu próprio caso claramente não necessito observar meu próprio comportamento para poder relatar acuradamente como me sinto. Apesar disso, insiste-se que, mesmo no caso da autoatribuição, as ligações com os critérios comportamentais para a atribuição de dor não foram perdidos. Se meu estado mental atribuído na ausência de critérios não é coerente com comportamento apropriado de minha parte, então errei em minha autoatribuição. Em outras palavras, mesmo que eu não precise fazer uso de critérios para atribuir dor a mim mesmo, sempre estarão disponíveis critérios na forma de meu comportamento público, contra os quais a correção ou erro de minha autoatribuição poderão ser verificados.

Examine a questão levantada no Exercício 9.5 antes de prosseguir a leitura.

### **EXERCÍCIO 9.5**

A explicação de Wittgenstein sobre como termos mentais são possíveis leva diretamente ao behaviorismo? Ela não se reduz, na verdade, à alegação de que

estar com dor realmente significa, afinal, comportar-se de uma determinada maneira?

É neste ponto que o assunto se torna bem difícil. A reação usual da parte de muitos filósofos neste caso consiste em focar no que significa dizer que o comportamento de uma pessoa é um critério para que ela esteja em certo estado mental. Como resposta, tem-se dito que um critério comportamental constitui uma parte essencial de um termo mental, mas o significado do termo não é exaurido pela descrição comportamental. Às vezes apresenta-se este ponto dizendo que o critério é uma evidência não-indutiva para a existência do estado mental. O propósito de chamar isso de evidência não-indutiva é indicar uma espécie mais profunda de ligação do que meramente a associação contingente entre comportamento e estados mentais, pressuposta pelo argumento por indução. Por outro lado, não há nenhuma relação direta de implicação lógica entre comportamento e estados mentais. Não se segue logicamente que, se uma pessoa está se comportando de uma determinada maneira, então ela deve encontrar se em um determinado estado mental. A noção de um critério, segundo a interpretação de Wittgenstein que estamos examinando, supostamente forneceria um meio-termo entre essas duas posições. Como diz Oswald Hanfling:

De acordo com [o argumento wittgensteiniano] [...] critérios podem constituir evidência que é mais forte do que mera indução [...] Por outro lado, esta evidência não é estrita à maneira de uma implicação dedutiva, de modo que o defensor dessa concepção pode admitir livremente – algo que seria absurdo negar – que pode haver comportamento de dor sem dor.<sup>17</sup>

Também nos é dito, pelo distinto comentador Peter Hacker, que “os critérios para se estar com dor não implicam em que a pessoa esteja com dor. Eles constituem uma boa evidência lógica, que é, em certas circunstâncias, revogável [isto é, que pode ser revogada]. Mas, se ela não é revogada, os critérios fornecem certeza”.<sup>18</sup>

Isso parece estar alegando que, normalmente, segue-se logicamente da satisfação dos critérios para se estar com dor que a dor esteja presente, mas podem existir circunstâncias especiais nas quais os critérios falham em permitir essa inferência. O defensor dos critérios, parece, deseja tomar de empréstimo os melhores traços de dedução e indução, mas deixa de lado o que provavelmente se mostrará embarçoso para uma explicação de como o conhecimento das outras mentes é possível.

Antes de continuar, faça uma pausa para examinar o Exercício 9.6.

### **EXERCÍCIO 9.6**

Quão inteligível lhe parece a ideia de um critério, como acabou de ser delineada? Discuta isso com outras pessoas antes de prosseguir a leitura.

A noção de um meio-termo, uma via media, entre dedução lógica estrita e indução não-lógica, fez alguns críticos se sentirem desconfortáveis. Parece que isso é uma maneira de querer guardar e comer seu doce ao mesmo tempo. Os críticos estão habituados a argumentar que não há espaço lógico para esse meio-termo. Dedução e indução, como formas de inferência, repartem o campo que está entre elas em sua totalidade, não deixando espaço para que uma terceira forma de inferência seja inserida. Mas assim o problema das outras mentes, como é tradicionalmente

concebido – isto é, como podemos passar do comportamento ao conhecimento do estado mental que lhe serve de suporte – volta a apresentar-se. Ficamos novamente com a escolha entre uma inferência deductiva do comportamento a partir de estados mentais e uma inferência indutiva, e nenhuma delas, como vimos, é adequada para fornecer o conhecimento que buscamos.

Também tem sido alegado, por exemplo, por Galen Strawson, que as concepções de Wittgenstein efetivamente levam ao behaviorismo analítico, afinal. Strawson afirma que, uma vez concedido que o comportamento deva desempenhar um papel crucial para possibilitar que os termos mentais adquiram seus significados, é difícil evitar o resvalo para o antirrealismo sobre o significado de termos mentais:

Uma vez que tenhamos aceitado o argumento dos teóricos W [wittgensteinianos], segundo o qual a referência a ocorrências publicamente observáveis participa essencialmente de qualquer explicação satisfatória do significado da palavra “dor”, pode ser difícil visualizar uma maneira de evitar o resvalo na concepção segundo a qual a referência a ocorrências publicamente observáveis é tudo que participa de qualquer explicação satisfatória do significado da palavra “dor”.<sup>19</sup>

A ideia básica por trás do antirrealismo é esta: o significado de um enunciado é dado por suas condições de assertibilidade, isto é, pelo conjunto de circunstâncias que, se satisfeitas, licenciariam a aplicação daquele enunciado àquelas circunstâncias, com o fim de produzir uma asserção verdadeira. Para o wittgensteiniano, as condições de assertibilidade que governam a aplicação, e, por conseguinte, o significado, de um enunciado sobre um estado mental consistem no comportamento em certas circunstâncias. Deve

estar claro que esta posição é essencialmente reducionista: os estados mentais são constituídos inteiramente pelo comportamento externo. Strawson criticou admiravelmente a posição wittgensteiniana ao defender que a redução de estados mentais a comportamento envolve um argumento em três passos. Partimos do modo como podemos dizer que alguém se encontra em um dado estado mental, chegando à posição segundo a qual o modo em que dizemos isso determina o que significamos quando atribuímos o estado em questão, e isso, por sua vez, fixa a natureza do estado. A epistemologia determina a semântica que, por sua vez, determina a ontologia. Isso é ilustrado na Figura 9.2.

Os wittgensteinianos dedicados negam, como vimos, que estejam comprometidos com um resvalo dessa espécie para o behaviorismo. Contudo, por razões de espaço e tempo, não poderei levar o debate adiante aqui, embora espere ter auxiliado na promoção de seu próprio pensamento a respeito desta difícil questão.

### **Termos mentais substituem o comportamento**

Que dizer da outra sugestão de Wittgenstein, a saber, que frases tais como “estou com dor” gradualmente passam a

substituir expressões naturais de dor? Um problema é que frases possuem propriedades que são diferentes das de expressões naturais de dor, de modo que é difícil ver como uma poderia ser a extensão ou o substituto da outra. Frases possuem valores de verdade; elas possuem tempos verbais; elas podem ser negadas; elas podem manter relações lógicas entre si; e elas também podem consistir em um componente de condicionais – por exemplo, se eu ficasse com dor, então eu deveria..., etc. Nenhuma dessas características se aplica ao caso de expressões naturais de dor, por exemplo, gemidos.

Um segundo problema tem a ver com o significado da afirmação de que o comportamento natural de dor, e as atribuições de dor em primeira pessoa que supostamente passam a substituí-lo, são expressões ou manifestações de dor. A maneira óbvia de se interpretar isso é dizer que minha contorção e meu gemido são manifestações de minha dor porque minha dor os suscita. Em outras palavras, um estado mental não-comportamental, privado, estar com dor, causa o comportamento externo de dor, assim como uma infecção, por exemplo, sarampo, causa o aparecimento de manchas na pele de uma pessoa. É usualmente por meio dessa manifestação externa do sarampo que

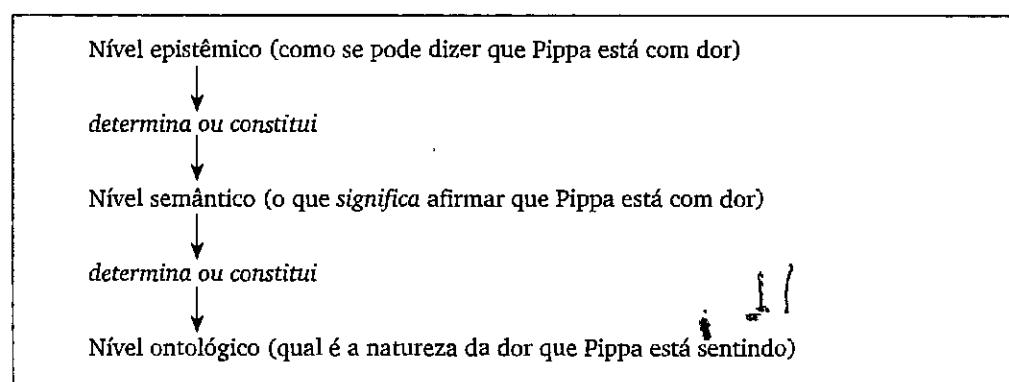


Figura 9.2 Neobehaviorismo.

detectamos pela primeira vez a presença do vírus. No caso de estados mentais, a confusão com esse tipo de explicação é que ela já pressupõe que estados mentais estejam presentes, e que eles manifestem sua presença no comportamento externo. Todavia, há uma enorme diferença entre estar com dor e estar com sarampo. Se estivermos incertos em relação a se as manchas na pele indicam sarampo, podemos fazer exames mais diretos para determinar se o vírus do sarampo está presente no corpo. Mas isso é precisamente o que estamos impossibilitados de fazer no caso dos estados mentais, exceto em nosso próprio caso. O cético sempre pode sustentar que o que aparecem ser manifestações naturais e adquiridas de estados mentais não são, na realidade, tais coisas. O comportamento parece indicar uma causa mental subjacente, mas na realidade não há tal causa, porque o que consideramos ser uma pessoa normal é na verdade um zumbi.

Há filósofos que sustentam que Wittgenstein não pode ser acusado de ser incapaz de fornecer uma solução para o problema das outras mentes, porque ele nunca tentou resolver esse problema para começo de conversa. Assim, Oswald Hanfling escreve:

Pensa-se às vezes que a solução de Wittgenstein, ou sua tentativa de fornecer uma solução, para o problema de outras mentes é uma de suas principais contribuições à filosofia. Contudo, seria difícil indicar um enunciado do problema em suas principais obras, sem falar em uma "solução" a ele por meio dos critérios. No argumento da linguagem privada, assim como em outros contextos das *Investigations*<sup>20</sup>, seu interesse principal é o significado, e não o conhecimento; o argumento é uma investigação sobre os limites da linguagem.<sup>20</sup>

Na medida em que Wittgenstein produziu uma solução para o problema de outras mentes, Hanfling sustenta que devemos buscá-la em sua observação de que "Se estivermos usando a palavra 'saber' do modo como ela é normalmente usada (e de que outro modo deveríamos usá-la), então outras pessoas muito frequentemente sabem que estou com dor".<sup>21</sup> Esta observação precisa ser entendida juntamente com outra alegação de Wittgenstein, de que não faz sentido, no meu próprio caso, dizer que sei que estou com dor. Isso porque em meu próprio caso não há lugar para dúvida ou erro, uma vez que meu argumento não está fundado sobre nenhuma evidência, e nenhum padrão de inferência entra em cena. A própria possibilidade de dúvida e engano é, desse modo, excluída, e por isso seria tentador dizer que, em seu próprio caso, no que toca a muitos tipos de estados mentais, não se pode deixar de conhecê-los. Mas, diz Wittgenstein, onde nada jamais pode ser considerado como fracasso cognitivo, então nada poderá ser considerado como sucesso. Seria melhor, então, dizer que, relativamente às próprias dores e outras sensações, não se aplicam nem o conceito de saber que alguém as possui, nem o conceito de fracassar em saber que alguém as possui. "Eu sei que estou com dor", consequentemente, não diz mais do que "Eu estou com dor"; as palavras "Eu sei" não adicionam nada de novo à frase.

No caso de outras pessoas, contudo, ainda que às vezes exista espaço para dúvida e erro concernente a qual seja seu estado mental, e, por conseguinte, à falta de conhecimento relativamente a ele, em outros casos não há fundamentos para dúvida e erro; e nesses posso dizer significativa e verdadeiramente que sei quais são os estados mentais dessas outras pessoas: "Eu posso estar tão certo a respeito

\* N. de T.: Trata-se das *Investigações Filosóficas*

das sensações de outra pessoa quanto de qualquer outro fato. [...] 'Mas se você está certo, isso quer dizer que você está fechando seus olhos à possibilidade de dúvida?' – Eles estão fechados".<sup>22</sup> Há casos em que não há lugar para dúvida. A dúvida, assim como o conhecimento, exige fundamentos, nos lembra Wittgenstein em *On Certainty*.<sup>23</sup> A mera possibilidade de que uma atribuição de sensação a outra pessoa possa vir a se mostrar falsa não pode constituir um fundamento para a dúvida. Uma pessoa em uma oficina, ou em um acidente de automóvel, está deitada no chão gritando, coberta de sangue, devido a terríveis ferimentos. Será que isso pode ter sido pré-arranjado, uma farsa? Haverá circunstâncias em que esta possibilidade simplesmente não poderá ser considerada, nas quais, segundo Wittgenstein, não se fecham os olhos à dúvida porque, devido à natureza do caso, eles já estão fechados: "Apenas tente – num caso real – duvidar do medo ou da dor de outra pessoa".<sup>24</sup>

Imagino, contudo, que essa resposta não será satisfatória para todo mundo. É claro que, na vida cotidiana, questões céticas sobre a crença na existência de um mundo externo, sobre a eficiência causal de estados mentais, e o problema da existência de outras mentes, não se colocam. Mas isso não significa que tais questões não possam ser enfatizadas filosoficamente. Ao tentar fornecer uma justificação racional para a indução, Hume admitiu que alguém pudesse dizer que suas práticas cotidianas contrariavam suas dúvidas. Em outras palavras, ele aceitou que não adotava uma atitude céтика em relação à indução, que se valia dela milhares de vezes por dia, sem jamais duvidar, por exemplo, de que o sol nasceria pela manhã, ou de que ele não pisaria no nada em vez do chão

sólido quando deixasse sua sala. Contudo, como Hume salientou, ainda que não haja problema algum com a observação de que ele não é um célico fora de seu gabinete, o que ele queria, se pudesse encontrar, era uma justificação racional para a indução.

Similarmente, na vida cotidiana, podemos pensar que é impossível duvidar de que alguém estivesse com dor, mas isso ainda não explica como poderíamos adquirir conhecimento dos estados mentais dos outros. Estados mentais, como foram concedido, não são idênticos ao comportamento. No caso de outras pessoas, temos conhecimento direto de seus comportamentos, mas não de seus estados mentais. Uma lacuna céтика se abre desse modo entre o conhecimento do comportamento de alguém e o conhecimento dos estados mentais que, supostamente, subjazem a esse comportamento. Uma vez que, aparentemente, nem o raciocínio dedutivo nem o indutivo podem preencher essa lacuna, e uma vez que sua eliminação, por meio da redução dos estados mentais ao comportamento, é inaceitável (ver Capítulo 4, seção "Ryle e as disposições"), aparentemente continuamos sem uma fundação sólida para rejeitar o ceticismo sobre as outras mentes.

### A tentativa de P. F. Strawson para solucionar o problema de outras mentes

Uma abordagem inspirada em Wittgenstein de como os termos mentais adquirem seu significado e uso, que contém muitas das características que distinguimos anteriormente, pode ser encontrada no trabalho de Peter Strawson, em seu livro *Individuals*.<sup>25</sup> Será instrutivo exami-

\* N. de T.: Em português: *Sobre a certeza*.

\*\* N. de T.: Em português: *Indivíduos*.

nar brevemente essa explicação antes de concluir este capítulo.

Uma passagem central de *Individuals* procede nas seguintes linhas: “É uma condição necessária para que alguém atribua estados de consciência, experiências, a si mesmo, do modo como alguém normalmente o faz, que esse alguém também possa atribuí-los, ou esteja preparado para atribuí-los, a outros que não ele mesmo”.<sup>25</sup> Em outras palavras, a autoatribuição só é possível porque a atribuição a outrem é possível. A autoatribuição só pode acontecer se a atribuição a outrem ocorrer. Da maneira como essa posição é apresentada, ela parece resolver o problema das outras mentes de um só golpe. A menos que eu possa atribuir estados mentais a outros, não poderei atribuir estados mentais a mim mesmo. Mas isso significa que não posso partir de meu próprio caso – isto é, atribuindo estados mentais a mim mesmo e então perguntando se eles podem ser atribuídos a outros. Apenas se eles já foram atribuídos a outros, ou, pelo menos, se eu estiver preparado a atribuí-los a outros, poderei formar o conceito de meus próprios estados mentais. Tendo argumentado em favor dessa posição, Strawson passa a explicitar como a atribuição a outrem é possível: “Alguém pode atribuir estados de consciência a si mesmo apenas se puder atribuí-los a outros. Alguém pode atribuí-los a outros apenas se puder identificar outros sujeitos de experiência. E não se pode identificar esses outros apenas como sujeitos de experiência, como possuidores de estados de consciência”.<sup>26</sup>

O pensamento aqui é que, se as pessoas forem concebidas como possuindo uma natureza puramente mental, ao modo das almas cartesianas, não será possível atribuir-lhes estados mentais. A razão é dupla. Em primeiro lugar, como vimos no Capítulo 2, se as pessoas forem concebi-

das como almas incorpóreas, então não seremos capazes de distinguir uma alma da outra. Mas desse modo não podemos formar o conceito de um sujeito individual de estados mentais. O princípio geral que subjaz a isso é que, para que um conceito tenha um significado, devemos poder especificar aquelas características de nossa experiência que, se satisfeitas, licenciariam a aplicação do conceito. Mas uma vez que não podemos ter nenhuma experiência de almas incorpóreas, não há como aplicar os conceitos correlacionados da mesma alma do ponto de vista numérico, e de uma alma diferente do ponto de vista numérico. Assim, nós na realidade não sabemos o que significa falar em uma alma ou em muitas almas (uma vez que muitas é uma coleção de umas).

Em segundo lugar, para que possamos atribuir características – estados mentais – a essas almas, devemos poder especificar aquelas condições experimentais que licenciariam a aplicação do estado mental em pauta. Mas, de novo, se esses estados são empiricamente indetectáveis, porque são de natureza não-física, então não podemos especificar as condições que, se satisfeitas, licenciariam a atribuição do estado mental particular em pauta. O que Strawson está implicando, e posteriormente explicitará, é que devemos conceber as pessoas como entidades materiais. Sua ocupação singular do espaço nos facilita distinguir entre dois sujeitos de experiência. Não importando o quanto duas ou mais pessoas assemelhem-se entre si, elas não podem ocupar o mesmo lugar no espaço ao mesmo tempo, de modo que sempre teremos uma base firme para distinguir uma pessoa de outra (ver Capítulo 2, seção “O problema de contar almas”). O fato de que pessoas possuem características físicas também torna possível a atribuição de estados mentais a elas. É o comportamento publicamente observável

dessas pessoas que fornece a base – ou, como diz Strawson, o critério logicamente adequado, um termo que é reminescente de Wittgenstein – para a atribuição de vários estados de consciência a elas. Assim, chegamos à mesma conclusão obtida no caso de nossa discussão sobre Wittgenstein: as atribuições em terceira pessoa tornam possíveis as atribuições em primeira pessoa e são logicamente indispensáveis para elas. Essa posição é claramente o inverso do cartesianismo. Para o cartesiano, alguém pode atribuir estados mentais a si mesmo puramente por estar privado e unicamente familiarizado com seus próprios estados mentais. Nenhuma familiaridade com, estados mentais de qualquer outra pessoa, ou o conhecimento deles, é pressuposto ou necessário para a auto-atribuição. Mesmo se não houver outros estados mentais, o cartesiano sustentará que ao menos você pode estar familiarizado com seus próprios estados mentais, e atribuí-los a si mesmo.

Quão bem-sucedido é Strawson em sua tentativa de neutralizar o problema das outras mentes dessa maneira? George Graham, em seu claríssimo livro *Philosophy of Mind: An Introduction\**, identifica duas dificuldades.<sup>27</sup> Em primeiro lugar, Graham pensa que, dado que a atribuição de estados mentais a outros é feita com base na observação de seus comportamentos, um defensor da teoria de Strawson, ou da teoria intimamente relacionada de Wittgenstein, terá dificuldades para evitar alguma forma de behaviorismo lógico. Já encontramos essa objeção anteriormente, do modo como fora formulada por Galen Strawson.

A segunda objeção de Graham é que o argumento de Peter Strawson incorre em petição de princípio: ele pressupõe que

deve ser verdadeiro que outras pessoas possuam mentes – isto é, que termos mentais possam ser verdadeiramente atribuídos a outras pessoas. Mas esse, obviamente, é o próprio ponto que está em questão.

O argumento que Graham apresenta foi provavelmente exposto de maneira mais clara por A. J. Ayer, em *The Concept of a Person\*\**.<sup>28</sup> A crítica central de Ayer contra o argumento de Strawson é que, mesmo que alguém tenha de fato atribuído, ou estivesse preparado para atribuir, ou mesmo tenha tido todas as razões para acreditar que estava justificado em atribuir, um estado mental a uma suposta outra pessoa, ainda assim não se seguiria que a atribuição fosse verdadeira, isto é, que a alegada outra pessoa realmente encontrava-se no estado mental atribuído, ou mesmo em qualquer estado mental em absoluto.

Para ilustrar o argumento que ele está apresentando, Ayer imagina o caso de uma criança mantida fora de qualquer contato com seres humanos, pelo menos em seus primeiros anos, que tenha sido criada por autômatos que têm semelhança e comportamento similares aos de seres humanos. Esses autômatos são construídos de tal maneira que, por exemplo, quando a criança bate neles, eles gritam ou revidam, e quando a criança lhes faz perguntas, eles respondem sacudindo e acenando com suas cabeças. A criança é instruída no uso da linguagem e em outras formas de comportamento por uma voz emitida por um alto-falante. Desse modo ela aprende os nomes dos objetos que estão no quarto em que ela é mantida, assim como o seu nome próprio, pronomes e demonstrativos. A voz também ensina à criança as palavras que descrevem seus próprios estados mentais, e ela

\* N. de T.: Em português: *Filosofia da mente, uma introdução*.

\*\* N. de T.: Em português: *O conceito de uma pessoa*.

também aprende como distinguir visão de imaginação, e memórias de eventos verídicos de memórias de sonhos. A voz enfatiza as similaridades entre os autômatos e a criança, sempre falando dos primeiros como se fossem, também eles, conscientes, e a criança acha que os diferentes estados de consciência atribuídos às máquinas ajustam-se a certos padrões gerais de comportamento que elas exibem, e também às mudanças em seus comportamentos. Desse modo a criança passa a atribuir estados mentais a si mesma e também aos autômatos. Em outras palavras, ela satisfaz a condição de Strawson, na medida em que está preparada para atribuir estados mentais a uma gama de outros seres similares, assim como a si mesma. Em resumo, ela adquire o conceito de uma pessoa e aprende a operar com ele. Claramente, contudo, todas as atribuições de estados mentais que ela faz, e está muito satisfeita em fazê-lo, aos autômatos são falsas, uma vez que os esses são meras máquinas sem mente. Essa é uma consequência importante do experimento mental. Como o próprio Ayer diz:

O exemplo não apenas mostra que alguém pode ser capaz de atribuir experiências a si mesmo, enquanto se engana invariavelmente ao atribuí-los aos outros, mas também que os critérios que são considerados logicamente adequados para atribuir experiências a outros não podem determinar nada além do fato de que algumas expressões estão corretas, que em tais e tais condições essa é a coisa certa a se dizer; não se segue necessariamente que o que é dito seja verdadeiro.<sup>29</sup>

Ayer continua:

Não há nenhuma garantia para assumirmos que o seu conceito de pessoa não é o mesmo que o nosso [...] Não se trata de ela possuir um conceito diferente do que é

ser consciente, ou de ela aplicar o conceito incorretamente, mas apenas de ela encontrar-se em uma situação na qual as coisas que ela tem toda a razão para pensar que são conscientes não o são realmente. Se ela fosse uma criança filósofa, poderia começar a questionar se seus companheiros realmente possuíam experiências do mesmo modo como ela possuía, e inferir que sim, a partir da semelhança percebida entre eles e ela mesma. Ou talvez, se ela fosse surpreendida por alguma qualidade estereotipada no comportamento deles, pudesse corretamente concluir o contrário. Seja qual for a conclusão à qual ela tenha chegado, seu ceticismo não seria sem sentido. Como ele poderia ser se fosse realmente justificado?<sup>30</sup>

Desse modo, parece que o argumento de Strawson não fornece uma via para superar o ceticismo sobre as outras mentes, assim como, infelizmente, tampouco a fornecem as tentativas aqui esboçadas de solucionar o problema. Esta é uma conclusão desapontadora, e tem-se a esperança de que seja equivocada. Ela possui, contudo, a aparência e a atmosfera de uma das questões perenes da filosofia, e creio que provavelmente continuará deixando perplexas as gerações futuras de pensadores.

### **Questões para reflexão**

1. Quais são os pontos fortes e fracos de se argumentar a partir do nosso próprio caso para a existência de outras mentes?
2. Se estados mentais fossem idênticos a estados cerebrais, então o problema de outras mentes seria idêntico ao problema de outros cérebros?
3. O conhecimento da própria mente depende do conhecimento de outras mentes?
4. Poderão os avanços científicos algum dia vir a possibilitar:

- (a) que sintamos a dor de outras pessoas?  
 (b) que vejamos diretamente os pensamentos de outras pessoas?
5. "Contentamento, aflição e divertimento não estão escondidos por trás do rosto que os manifesta, mas nele encontram-se visíveis. O que nós tão enganosamente chamamos de 'o interno' injeta-se no externo [...] Vemos amizade ou animosidade em um rosto e não inferimos sua presença a partir da disposição dos músculos faciais (que sequer poderíamos descrever)." (HACKER, Peter. *Wittgenstein*. London, Phoenix, 1997, p. 43. Como a observação de Hacker deve ser interpretada? Você concorda com ela?)
6. O que significa falar na descrição do comportamento como um critério para a existência de estados mentais? O quanto convincente lhe parece a explicação criterial dos significados dos termos mentais? Ela leva, em última instância, ao neobehaviorismo?
- KENNY, A. *Wittgenstein*. London, Allen Lane, The Penguin Press, 1973. Se você deseja obter uma compreensão rápida do pensamento de Wittgenstein, incluindo especialmente o argumento da linguagem privada, leia esse livro.
  - HACKER, P. *Insight and Illusion*. Oxford, Clarendon Press, 1972, onde se encontra uma das melhores introduções à filosofia da mente de Wittgenstein.
  - HACKER, P. *Wittgenstein*. London, Phoenix, 1997 é curto e incisivo e contém observações interessantes e provocativas sobre o problema das outras mentes.
  - BUDD, M. *Wittgenstein's Philosophy of Psychology*. London, Routledge, 1989 apresenta uma investigação sistemática das concepções de Wittgenstein sobre a natureza da mente. Trata-se de uma leitura essencial.

### As seguintes obras também merecem ser lidas

#### Sugestões para leituras aprofundadas

- LOCKE, D. *Myself and Others; A Study in Our Knowledge of Other Minds*. Oxford, Clarendon Press, 1968 é extremamente útil caso você o encontre, mas o livro pode estar esgotado.
- MCGINN, M. *Wittgenstein*. London, Routledge, 1997 é útil especialmente para a discussão sobre a possibilidade de uma linguagem logicamente privada e sobre a viabilidade da distinção entre o "interno" e o "externo".
- MCGINN, C. *Wittgenstein on Meaning*. Oxford, Blackwell, 1984; GRAHAM G. *Philosophy of Mind: An Introduction*. Oxford, Blackwell, 1993; AYER, A. J. *The Concept of a Person*. London, MacMillan, 1963; BILGRAMI, A. "Other Minds", in: DANCY, J. e SOSA, E., eds. *The Blackwell Companion to Epistemology*. Oxford, Blackwell, 1992.
- AVRAMIDES, A. *Other Minds*. London and New York, Routledge, 2001. Escrito durante dez anos, o livro oferece uma discussão admiravelmente lúcida e completa do assunto.

## NOTAS

1. Obviamente poderia haver uma dúvida sobre a existência de outros cérebros, caso invocássemos o ceticismo sobre a existência de um mundo externo. Mas essa dúvida abrange igualmente a existência de meu próprio cérebro e de meu próprio corpo, e não meramente daqueles de outras pessoas. Eu deixo esse ceticismo generalizado, extremo, de lado, porque ele não possui relevância ou aplicação ao problema específico da existência de outras mentes.
2. Excluo a possibilidade de telepatia por duas razões. Em primeiro lugar, aparentemente sua existência jamais foi convincentemente estabelecida. Em segundo lugar, e de maneira mais importante, está longe de ser claro que essa noção seja coerente. Como, por exemplo, se supõe que eu poderia identificar certos pensamentos que chegam à minha mente como sendo seus, em vez de meus?
3. Estritamente falando, o comportamento em pauta deverá ser caracterizado em termos não-intencionais, isto é, em termos que não atribuem já de saída ou que não implicam a existência de estados mentais da alegada pessoa com a qual estamos envolvidos. Se não formos cuidadosos na caracterização do comportamento em termos neutros, de cunho não-mental, a questão sobre a existência de outros estados mentais será circularmente decidida em favor dessa existência. Devemos evitar pressupor já de saída a existência daquilo cuja existência deveríamos estabelecer.
4. PUTNAM, H. "Brains and Behaviour", in: *Mind, Language and Reality*, vol. II. Cambridge, CUP, 1975, p. 330 (itálicos no original).
5. Isso levou alguns filósofos a questionar se a inferência indutiva é realmente um tipo de inferência, e originou a notória dificuldade de como justificar a indução, uma questão complexa por si mesma, que não pode ser levada adiante aqui. Ver HUME, David. *Uma Investigação Sobre o Entendimento Humano*, seção IV, partes I e II, e seção V parte I, para uma famosa discussão sobre o problema da indução.
6. MILL, J. S. *An Examination of Sir William Hamilton's Philosophy*, 6. ed. New York, Longman's Green and Co., Inc., 1889, p. 243.
7. Agradeço a Roger Lindsay por esta analogia.
8. KANT, I. *Critique of Pure Reason*, [SMITH, N. Kemp, trad.] London, Macmillan, 1963, A52, p. 93.
9. FREGE, G. "On Sense and Meaning", in: GEACH, P. e BLACK, M. eds., *Translations from the Philosophical Writings of Gottlob Frege*, 3. ed., Blackwell, Oxford, 1980, p. 59. Ver também GUTTEPLAN, S. in: *A Companion to the Philosophy of Mind*. Oxford, Blackwell, 1994, p. 55.
10. WITTGENSTEIN, L. *The Blue and Brown Books*. Oxford, Blackwell, 1964, p. 2.
11. WITTGENSTEIN, L. *Philosophical Investigations*. Oxford, Blackwell, 1968, parágrafo 33.
12. Ibid.
13. WITTGENSTEIN, L. *Zettel*. Oxford, Blackwell, 1967, parágrafo 198.
14. WITTGENSTEIN, L. *The Blue and Brown Books*, p. 5.
15. WITTGENSTEIN, L. *Philosophical Investigations*, parágrafo 198.
16. Ibid., parágrafo 580.
17. HANFLING, O. *Wittgenstein's Later Philosophy*. London, Macmillan, 1989, p. 122. Relacionado a esse assunto, ver também HACKER, P. Wittgenstein. London, Phoenix, 1997, p. 19.
18. HACKER, P. *Wittgenstein*, p. 38.
19. STRAWSON, G. *Mental Reality*. Cambridge, MA, MIT Press, 1994, p. 223-4.
20. HANFLING, O. *Wittgenstein's Later Philosophy*, p. 126.
21. WITTGENSTEIN, L. *Philosophical Investigations*, parágrafo 246.
22. WITTGENSTEIN, L. *Philosophical Investigations*, parágrafo 224e.
23. WITTGENSTEIN, L. *On Certainty*. Oxford, Blackwell, 1969.
24. WITTGENSTEIN, L. *Philosophical Investigations*, parágrafo 303.
25. STRAWSON, P. F. *Individuals*. London, Methuen, 1959, p. 99.
26. Ibid., p. 100.
27. GRAHAM, G. *Philosophy of Mind: An Introduction*. Oxford, Blackwell, 1993, p. 39-40.
28. AYER, A. J. *The Concept of a Person*. London, Macmillan, 1963, p. 105-6.
29. Ibid., p. 108 (ênfase minha).
30. Ibid.

# 10

## IDENTIDADE PESSOAL COMO CONTINUIDADE FÍSICA

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender a diferença entre identidade qualitativa e numérica.
- Entender a diferença entre como dizemos que uma pessoa do presente é numericamente idêntica a uma pessoa do passado, e o que é constitutivo da identidade da pessoa ao longo do tempo.
- Entender a diferença entre a teoria reducionista e a não-reducionista da identidade pessoal.
- Entender a teoria da continuidade física e como uma de suas versões, a teoria da continuidade corporal, pode ser atacada usando o experimento mental dos transplantes de cérebro.
- Entender que a continuidade do cérebro é mais plausível que a continuidade do corpo como uma condição necessária da identidade pessoal.
- Entender como o experimento mental da fissão do cérebro leva à visão de que a sobrevivência sem identidade é possível.

### Introdução

Os dois últimos capítulos deste livro são dedicados ao tema da identidade pessoal. Este capítulo introduz o problema do que constitui a identidade pessoal e foca na continuidade física como uma abordagem dele. O Capítulo 11 examina a afirmação contrastante de que a continuidade psicológica é o mais importante para nossa identidade ao longo do tempo,

equilibrando esta abordagem com uma retrospectiva do local e da importância da continuidade física na obtenção de uma teoria satisfatória da identidade pessoal.

O problema da identidade pessoal pode ser entabulado inicialmente desta maneira. O que faz com que seja a pessoa presente – vamos chamá-la de Shaun – no tempo  $t_1$  uma e a mesma pessoa que Shaun em um tempo anterior,  $t_0$ ? Shaun, vamos supor, nasceu aproximadamente

19 anos atrás. Desde então, mudou consideravelmente a partir dos choramingos infantis e vômitos nos braços de sua mãe. Ele nasceu praticamente sem cabelo, mas agora tem uma bela cabeleira, preta e lisa. Suas unhas então eram curtas, sem pintura e largas. Agora elas são longas, finas e pintadas de preto (segundo a última moda). Em certo sentido, Shaun não é mais a pessoa que ele fora. Suas características mudaram consideravelmente ao longo do período de 19 anos. Ainda assim, Shaun em  $t_1$  e Shaun em  $t_0$  são uma e a mesma pessoa. Então Shaun, pode parecer, tanto é, como não é, a mesma pessoa. A contradição expressa nessa última sentença é, porém, apenas aparente. Isso assim é porque a sentença possui duas noções bem diferentes de identidade e igualdade. Ainda assim, não há colisão no significado quando afirmamos que Shaun tanto é, como não é, a mesma pessoa.

Filósofos marcam a distinção entre esses dois sentidos de “mesmo”, dizendo que, qualitativamente, Shaun não é a mesma pessoa que ele fora. Porém, numericamente, Shaun em  $t_1$  é idêntico, um e o mesmo, que Shaun em  $t_0$ . Para ter certeza que a distinção entre identidade numérica e qualitativa é clara, examine os seguintes exemplos.

**Identidade qualitativa** – Se uma coisa é qualitativamente idêntica a outra coisa, então ela se parece exatamente com a outra coisa. Por exemplo, seu carro é exatamente igual ao meu. Você possui um Ford Fiesta, e eu também. Seu carro se parece exatamente ao meu e tem todas as mesmas características, por exemplo, sistema de injeção de combustível, janelas elétricas, e assim por diante. Se seu carro fosse destruído, o meu ainda existiria, e vice-versa. Gêmeos idênticos também fornecem um exemplo pronto: Felipe e Pedro são tão semelhantes que frequentemente não posso dizer a qual estou ensinando

sem olhar para a lista de chamada (felizmente eles estão em turmas diferentes!).

**Identidade numérica** – Se uma coisa é numericamente idêntica a outra coisa, então de fato temos uma coisa em nossas mãos, e não duas coisas. Temos um e o mesmo Ford Fiesta. Saio com ele nas segundas, quartas e sextas. Você, nas terças, quintas e sábados. Nós dois saímos com ele aos domingos. Se o carro for destruído, nós dois simultaneamente ficamos sem carro. Outros casos ilustram o ponto. Os dorminhocos, como você lembrará do Capítulo 3, seção “Evitando objeções iniciais à teoria da identidade”, pensavam que a estrela matutina e a estrela vespertina são dois corpos celestiais bem diferentes. Contudo, elas são um e o mesmo objeto, o planeta Vênus. Édipo Rei casou com sua mãe, mas ele não fez isso consciente e intencionalmente, porque não sabia que a Rainha Jocasta era a uma e a mesma pessoa que sua mãe. A polícia não suspeitava, quando capturou Mr. Hyde, que de fato estivesse prendendo Dr. Jekyll, que estava terrivelmente transformado.

Em relação ao tópico da identidade pessoal, o interesse dos filósofos é na identidade numérica, e não na qualitativa. Então, para retornar à nossa questão inicial, em que consiste a identidade numérica de uma pessoa ao longo do tempo? O que então torna Shaun agora uma e a mesma pessoa que ele fora então?

### A metafísica e a epistemologia da identidade pessoal

Antes de prosseguirmos, porém, mais uma distinção precisa ser feita. Perguntando o que torna o Shaun agora numericamente idêntico ao Shaun de outrora, não estou meramente perguntando como digo que Shaun é uma e a mesma pessoa

a quem ensinei na semana passada. Em uma situação normal, posso dizer que o estudante ao qual estou ensinando agora é uma e a mesma pessoa que enviei à biblioteca na semana passada, tendo por base sua aparência física, suas características corporais, os tipos de roupas que ele usa habitualmente e as inconfundíveis forma e cor de suas unhas. Mas, evidentemente, nenhuma dessas características é constitutiva da identidade de Shaun. Isso significa dizer que nenhuma dessas características é essencial para que o Shaun de hoje seja uma e a mesma pessoa que o Shaun de ontem.

O homem careca, gordo, com barba ruiva, unhas foscas e de meia idade que estará frente a mim daqui a 25 anos também pode ser o magro, com cabelo preto, sem barba, com unhas pretas Shaun a quem eu ensinei muitos anos antes. E, possivelmente, mesmo se eu colhesse as impressões digitais de Shaun enquanto ele era um estudante e as comparasse com as de meu visitante de 44 anos, 25 anos, depois, descobrindo que elas não combinam, isso não seria suficiente para provar que o visitante não é Shaun. Talvez, com algum propósito nefasto em mente, ele tivesse feito alterar cirurgicamente suas impressões digitais. Talvez, ainda, por meio de algum processo biológico natural, suas impressões digitais tenham sofrido alterações, assim como ocorreu com o restante de sua aparência. Afinal de contas, é um fato meramente contingente, e não necessidade lógica, que as impressões digitais das pessoas não se modifiquem significativamente ao longo do tempo, e também que as impressões digitais de dois indivíduos nunca sejam qualitativamente exatamente as mesmas. Se ocorresse que as impressões digitais de todos fossem qualitativamente idênticas, o que logicamente poderia ser o caso, então as impressões digitais poderiam não servir como um meio

útil, mas falível, de estabelecer a identidade. Teríamos de recorrer a algum outro critério – DNA, por exemplo. Como, na prática, dizemos se uma pessoa ao tempo  $t_1$  é ou não uma e a mesma pessoa que alguém no tempo  $t_0$ , não pode dar uma resposta, então, quanto ao que consiste a identidade de uma pessoa, o que é essencial, lógica ou conceitualmente, para a identidade da pessoa e não somente e contingentemente um signo conveniente disto.

O que faz Shaun agora ser uma e a mesma pessoa que Shaun de outrora é uma referência metafísica sobre a natureza da identidade de pessoas, o que é logicamente indispensável para isso, e não apenas uma questão epistemológica similar a como dizermos que a pessoa que encontramos agora é uma e a mesma que a pessoa que encontramos anteriormente. A evidência falível à qual apelamos, tentando determinar questões sobre a identidade pessoal, não nos diz o que é logicamente constitutivo da identidade de pessoas ao longo do tempo. Respondendo à questão, precisamos explicar detalhadamente as condições necessárias e suficientes para que uma pessoa em  $t_1$  seja uma e a mesma que uma pessoa em  $t_0$ . Em outras palavras, precisamos completar a fórmula: "Shaun em  $t_1$  é uma e a mesma pessoa que Shaun em  $t_0$ , se e somente se..."

A noção de condições necessárias e suficientes pode ser esclarecida por meio de alguns exemplos simples. Uma condição necessária para algo ser um cavalo é que ele tenha sangue quente, mas simplesmente isso não é suficiente, pois muitos animais de sangue quente não são cavalos – pássaros, por exemplo. Uma condição pode ser suficiente, mas não necessária. Por exemplo, é suficiente para ser britânico que você tenha nascido na Grã-Bretanha, de pais britânicos. Mas isso não é necessário, pois a cidadania

britânica pode ser adquirida pela naturalização ou pelo casamento. Com o que, então, se pareceria uma afirmação das condições necessárias e suficientes para a nacionalidade britânica? Bem, talvez algo como isto: uma pessoa é uma cidadã britânica se, e somente se, ela nasceu de pais britânicos na Grã-Bretanha, ou ela adquiriu a nacionalidade britânica por meio da naturalização, ou casou com alguém britânico, ou... ou... (Não sei precisamente quais seriam essas outras condições). A satisfação de qualquer uma dessas condições é suficiente para tornar alguém britânico, mas nenhuma, tomada singularmente, é necessária, pois a satisfação de uma das outras condições teria, igualmente, cumprido sua função. Porém, a lista de alternativas, considerada como um todo, é necessária, pois ao menos uma das várias alternativas tem de ser satisfeita. Não importa qual, mas ela deve estar nessa lista de alternativas e não fora dela. Nesse sentido, a lista de alternativas – o que os filósofos chamam de uma disjunção – é uma condição necessária (ver Capítulo 3, seção “A identidade espécime-espécime e a impossibilidade de redução”).

### **Reducionismo e não-reducionismo**

O pressuposto da seção anterior foi de que há algo que é constitutivo da identidade pessoal – isto é, há uma maneira de expressar isso de outra forma, que não faz uso do próprio conceito de identidade pessoal, de que ela consiste. Fazer essa suposição é abraçar uma forma de reducionismo, vincular-se à crença de que o conceito de identidade pessoal pode ser expresso por meio de outros conceitos sem perda de significado. Você se lembrará de que um movimento análogo foi feito

pelos behavioristas analíticos (Capítulo 4) quando alegaram que tudo que queremos dizer sobre a mente pode ser captado por falar sobre comportamentos reais ou possíveis.

Porém, há filósofos que discordam que o conceito de identidade pessoal é analisável da maneira proposta pelos reducionistas. Esses filósofos – não-reducionistas – sustentam que o conceito de identidade pessoal é primitivo e não analisável, subscrevendo o dito do Bispo Butler de que “Tudo é o que é, e não alguma outra coisa”. Um grande problema para o behaviorismo analítico era que ele envolvia circularidade: qualquer tentativa de traduzir o falar sobre estados mentais em falar sobre o comportamento terminava tendo de introduzir na análise a própria noção de estado mental que ela tinha o propósito de analisar. Similarmente, também os não-reducionistas, no que tange à identidade pessoal, alegam que qualquer análise pretendida do conceito irá inevitavelmente, em alguma medida, precisar recorrer à noção não-analisável de uma e a mesma pessoa que está procurando analisar. Posteriormente, retornaremos ao não-reducionismo em relação à identidade pessoal.

### **Teorias reducionistas da identidade pessoal**

Ao longo deste livro, temos visto que, contrastando com objetos físicos ordinários tais como rochas, árvores e artefatos, os seres humanos podem ser vistos tanto da perspectiva da primeira quanto da terceira pessoa. Da terceira pessoa, ponto de vista objetivo, os seres humanos são organismos físicos vivos cujas operações internas e comportamento externo estão abertos a escrutínio e observação públicos. Mas as pessoas também têm uma

perspectiva do mundo privada e subjetiva, que não parece ser acessível aos outros. Esse contraste epistêmico entre as perspectivas da primeira e da terceira pessoas também encontra sua expressão nas diferentes maneiras pelas quais fazemos julgamentos sobre a identidade dos outros, por comparação com os julgamentos que fazemos sobre nossa própria identidade.

Da perspectiva da terceira pessoa, sei que Shaun é uma e a mesma pessoa a quem ensinei na semana passada, porque tenho uma boa razão para acreditar que o ser humano vivo de carne e ossos na minha frente agora é fisicamente contínuo com a pessoa que outrora ensinei. Poderei, naturalmente, estar equivocado. Um sósia de Shaun pode tê-lo substituído, de tal maneira que, se eu retroativamente traçasse o percurso do sósia ao longo de espaço e tempo, descobriria que aquele não coincide com a trajetória de espaço e tempo traçada pelo Shaun original. Porém, não apenas a substituição de um sósia é extremamente improvável, como também, se eu tivesse alguma dúvida de que tal peça pudesse ter sido pregada, poderia sempre assegurar-me de que nunca deixei Shaun fora de minha vista desde a aula da última semana até a desta semana. Na prática isso seria muito difícil, mas seria, em princípio, realizável. Em um caso como esse, como digo que o Shaun de agora é um e o mesmo que o Shaun de outrora, e o que significa dizer que o Shaun de agora é um e o mesmo que o Shaun naquele tempo parecem coincidir. A identidade de Shaun ao longo do tempo consiste na continuidade física de um corpo humano vivo através de espaço e tempo. Essa é a resposta do senso comum, que retornaria à questão do que constitui a identidade pessoal, e tanto quanto possível parece ser correta.

Mas e sobre meu próprio caso? Antes de prosseguir a leitura, pense sobre isso e sobre a questão posta no Exercício 10.1.

### **EXERCÍCIO 10.1**

Como sei que de fato fui eu que ensinei Shaun na última semana, e que não era outro professor que me substituiu e deu a aula?

A resposta breve é que me lembro de estar aqui ensinando Shaun na semana passada. Tenho boa memória e lembro distintamente de ter vindo a esta sala na semana passada e ensinado Shaun. Em um caso como esse, tenho acesso ao conhecimento da identidade da pessoa que ensinou Shaun – a saber, eu mesmo – de uma perspectiva de primeira pessoa, não-observacional. Recordo a experiência de estar aqui na sala de aula ensinando Shaun, e é assim que sei que de fato era eu mesmo, uma e a mesma pessoa que se lembra agora que estava na sala de aula há uma semana. A questão “lembro-me distintamente de ensinar Shaun na semana passada, mas fui eu que o ensinei naquela vez?” é sem sentido. Se realmente me lembro, então devo agora ser uma e a mesma pessoa que ensinou Shaun naquele momento.

A situação é totalmente diferente daquela em que uma pessoa tomou o meu lugar por uma semana. Em um caso como esse, ao procurar na sala dos professores, poderia concebivelmente equivocar-me quanto à identificação do professor que assumiu minhas aulas, pensando equivocadamente que era a pessoa no canto, enquanto na realidade era a pessoa à porta. É claro, posso ter estado sob alguma espécie de delírio quando afirmei que ensinei Shaun na última semana. Em primeiro lugar, posso ter dado alguma aula, mas a pessoa que ensinei pode ter sido

o sócio de Shaun, mencionado anteriormente. Ou, mais radicalmente, poderia ser o caso de eu, possivelmente, não ter ensinado Shaun na última semana porque estava em casa com gripe, e a doença causou alguma estranha alteração da memória ou amnésia. Todavia, a possibilidade da ocorrência dessas memórias equivocadas não altera o fato de que, quando me lembro, posso efetivamente saber "de dentro", isto é, de uma perspectiva da primeira pessoa e não-observacionalmente, que agora sou uma e a mesma pessoa que a pessoa que ensinou Shaun na semana passada. Outras pessoas não podem saber dessa maneira. Sua única maneira de saber isso é pela observação do ser humano vivo, Keith Maslin, que existiu física e continuamente da semana passada até esta, traçando um percurso publicamente verificável e sem rupturas, através de tempo e espaço.

Estamos agora em posição para apreciar que essas duas maneiras pelas quais a identidade de uma pessoa pode ser conhecida deram origem a duas importantes teorias do que constitui a identidade pessoal. A primeira dessas é a teoria da continuidade física, a qual já referi em parte. Uma pessoa é agora uma e a mesma que uma pessoa do passado porque a pessoa presente é fisicamente contínua com a pessoa do passado. O filósofo Bernard Williams, em um artigo escrito em 1956-1957, faz uso do critério de identidade da continuidade física.<sup>1</sup> Ele imagina o caso em que dois homens, Charles e seu irmão Robert, fazem afirmações sobre memórias que são muito semelhantes às que sabemos da vida de Guy Fawkes, assim como exibem traços de personalidade similares aos de Fawkes. Devemos concluir que Fawkes retornou de alguma maneira à vida na pessoa de um desses irmãos? Se fizermos isso, teremos de tentar determinar qual dos homens é Fawkes. Pense

cuidadosamente sobre isso e examine o exercício 10.2 antes de continuar.

### **EXERCÍCIO 10.2**

Como vamos determinar qual pessoa era Fawkes se é que alguma delas era?

A resposta seria fornecida, afirma Williams, se descobrissemos qual dos dois irmãos, Charles ou Robert, é fisicamente contínuo com Fawkes. Apenas a pessoa que segue a mesma trajetória contínua ao longo de tempo e espaço que Fawkes pode ser identificada com ele.

A segunda teoria é a teoria da continuidade psicológica. De acordo com ela, o que é importante para a identidade de uma pessoa ao longo do tempo não é a identidade do corpo de uma pessoa ou, de fato, qualquer parte física de seu corpo, tal como o cérebro, mas a continuidade de sua vida mental: a continuidade de memória, traços de caráter e habilidades intelectuais e artísticas. Imagine uma pessoa cujas memórias e traços de caráter, assim como suas habilidades, até mesmo ser capaz de comer, andar e falar, são completamente apagadas por um violento remanejamento de seu cérebro. De acordo com a teoria da continuidade física, teríamos uma e a mesma pessoa em nossas mãos depois do apagamento como tivemos antes da destruição de todas as suas memórias, traços de caráter e habilidades. Isso porque, fisicamente, o corpo e o cérebro da pessoa após a drástica cirurgia cerebral são os mesmos que o corpo e o cérebro antes da operação. Mas precisamos perguntar: seria realmente uma e a mesma pessoa, ao contrário do que a teoria da continuidade física sustenta? Se, após o apagamento, a pessoa gradualmente adquiriu as habilidades de andar e falar, desenvolveu novos traços de caráter

totalmente diferentes dos antes possuídos e começou a guardar todo um novo conjunto de memórias, seria realmente uma e a mesma pessoa agora que aquela antes da cirurgia cerebral, apesar de ter um e o mesmo corpo e cérebro?

Considerações como essas têm levado alguns filósofos, a partir de John Locke no século XVII, a sustentar que a identidade pessoal possa ser analisada em termos simplesmente de continuidade e conectividade da vida mental de uma pessoa, e que, enfim, a identidade corporal e a continuidade física são irrelevantes. Lembre-se de que as teorias da continuidade física e da continuidade psicológica, apesar de suas diferenças óbvias e radicais, pertencem ambas ao campo reducionista, uma sustentando que a identidade pessoal consiste em certos fatos mais particulares sobre a identidade corporal ou física, a outra afirmando que a identidade pessoal não deve ser explicada em termos de memória e outras relações psicológicas. Agora examinaremos a primeira dessas teorias, antes de nos voltarmos à teoria da continuidade psicológica, no Capítulo 11.

### A teoria da continuidade física

Resumidamente, essa teoria sustenta que uma pessoa do presente é uma e a mesma que uma pessoa do passado se, e somente se, a pessoa em questão tem um e o mesmo corpo hoje que tinha ontem. Com mais detalhes, o que isso significa é que o corpo de Shaun precisa ter existido continuamente ao longo do período de tempo em questão. Em outras palavras, ele precisa reunir o mesmo critério para a identidade ao longo do tempo, assim como os objetos físicos ordinários, a saber, a continuidade espaço-temporal – a capacidade de traçar uma única trajetória

contínua ao longo de tempo e espaço. Mas aqui precisamos ser cuidadosos. Suponha que Shaun foi submetido a uma cirurgia radical na qual seus braços e pernas foram separados e então costurados de volta. O corpo de Shaun é um e o mesmo depois da separação e reunião dos membros? Certamente pareceria que sim. Considere outra possibilidade. Shaun é rapidamente congelado e cortado em cubos. Após um tempo, esses cubos são reunidos novamente e então descongelados. Shaun deserta da operação aparentemente tranquilo. O corpo de Shaun é um e o mesmo que era antes do esquartejamento?

Antes de prosseguir a leitura, veja o Exercício 10.3 e examine sua reação à questão apresentada.

### EXERCÍCIO 10.3

A identidade numérica de Shaun é preservada, apesar do esquartejamento e da reunião das partes? Que argumentos você pode criar para defender e para atacar essa afirmação?

Acredito que há uma razão para supor que a identidade do corpo de Shaun é preservada. Caso alguém duvide disso, considere a situação em que quero ir andar de bicicleta em alguma parte distante do país, mas não quero chegar a essa região distante tendo de ir de bicicleta até lá. Como minha bicicleta não cabe em meu carro, desmontei-a, coloquei-a no porta-malas e remontei-a quando alcancei minha destinação. A bicicleta remontada é uma e a mesma que a bicicleta desmontada? Seria loucura supor que não. A separação e a reunião do corpo de Shaun em seu estado congelado, após o qual ele é descongelado e volta a funcionar normalmente, é, sustento, em princípio não diferente da separação e reunião da bicicleta.



A identidade do corpo de Shaun ao longo do tempo também não é perturbada pelo fato de que o corpo está continuamente em renovação. Na medida em que o tempo passa, Shaun solta células mortas da pele e perde várias partículas das partes internas na medida em que seus órgãos envelhecem. De fato, os cientistas nos dizem que, ao longo de um período de tempo, todas as moléculas do corpo de Shaun mudam. Ainda assim, como a mudança é gradual, um e o mesmo corpo continua a existir. Compare com o caso do navio de Teseu no famoso enigma da Grécia antiga a respeito da identidade.

Teseu tem um navio de carvalho composto de mil tábuas de carvalho em boas condições, mas ele viu um navio de teca e desejou ter um desses. Por conseguinte, ele removeu uma tábua de carvalho de cada vez, substituindo-a por uma tábua de teca. Leva um minuto para remover uma tábua de carvalho e substituí-la por uma tábua de teca; então, após mil minutos, o navio de Teseu é composto inteiramente de teca. Certamente ele sofreu grande mudança qualitativa. Ele era inteiramente feito de carvalho, e agora, inteiramente de teca. Mas se pode argumentar que ele numericamente ainda seja um e o mesmo navio, porque ao longo das mudanças na matéria que o compõe, sua forma – isto é, sua estrutura organizacional – e a função à qual a estrutura serve são continuamente preservadas. A identidade de muitas coisas é assim: árvores, edifícios, clubes e organizações em geral, ainda que suas partes constituintes mudem gradualmente, não perdem suas identidades numéricas. Não obstante a mudança completa nos estudantes, no corpo docente e concebivelmente em todos os edifícios, o Esher College do futuro será um e o mesmo que o Esher College de hoje, por causa da preservação contínua da orga-

nização e da função da faculdade, não obstante suas alterações radicais, mas graduais.

Existem algumas mudanças, porém, que pareceriam destruir a identidade numérica de Shaun. Considere o seguinte cenário. Shaun entra em uma máquina de desintegração molecular. Seu corpo é partido em suas moléculas constituintes, que são então espalhadas aos quatro ventos. Posteriormente, por puro acaso, um corpo que é qualitativamente idêntico ao corpo de Shaun no momento da dispersão é reconstituído a partir de suas moléculas originais, indo os tipos corretos de moléculas exatamente para os mesmos tipos de localização que ocupavam antes da separação. Seria esse novo corpo um e o mesmo que o antigo corpo que foi destruído? Parece que não. Quando as moléculas são espalhadas, elas seguem existindo como coisas individuais e conservam suas identidades, mas a identidade do corpo que elas costumavam constituir está perdida. Isso porque a estrutura organizacional, que é essencial para Shaun ser a entidade individual ou substância lógica que é, não é preservada quando seu corpo é decomposto nos microelementos que o compreendem.

Parece ser um princípio metafísico sólido que uma substância lógica não pode ter dois inícios. Consequentemente, mesmo após todas as moléculas terem sido reunidas e colocadas em ordem em um corpo que é qualitativamente idêntico ao antigo corpo de Shaun, este novo corpo ainda não será o antigo corpo de Shaun, apesar de ser feito da mesma matéria que constituía o corpo original de Shaun e apesar de ser qualitativamente indistinguível deste.

Suponha, menos radicalmente, que as moléculas biológicas de Shaun são substituídas por moléculas inorgânicas que, após serem gradualmente absorvidas nos vários órgãos, fossem capazes de exe-

cutar exatamente as mesmas funções que foram executadas pelos órgãos originais, antes da substituição gradual de suas partes. Em outras palavras, o corpo de Shaun termina sendo algo como o corpo de um androide. É um e o mesmo corpo após essas mudanças? Isso é muito mais difícil de solucionar. O fato de que as mesmas funções são executadas nos inclina a dizer que é um e o mesmo corpo. Por outro lado, podemos significativamente ainda falar de um corpo humano, uma vez que ele consiste inteiramente de partes inorgânicas? Mudanças radicais desse tipo, pode-se dizer, efetivamente destroem a identidade do corpo. Deixo ao leitor tomar sua decisão quanto a isso.

### **Continuidade física e transplantes cerebrais**

Mesmo se alguns tipos de continuidade física devessem vir a ser uma condição necessária para a identidade pessoal, ela claramente não é suficiente para isso. Se Shaun encontra-se morto no chão à minha frente, então, considerado como uma massa de matéria física organizada, Shaun morto tem um e o mesmo corpo que ele tinha quando estava vivo. Mas a existência do corpo morto de Shaun não pode fornecer a identidade da pessoa entre ontem e agora, porque a pessoa não está mais aqui. Ainda que possamos falar da pessoa morta, estritamente não temos mais uma pessoa em nossas mãos, mas uma ex-pessoa, um cadáver, o que restou do que era antes uma pessoa. Para que a continuidade do corpo de Shaun possa inclusiva ter a possibilidade de ser necessária para a identidade de Shaun ao longo do tempo, parece que precisamos especificar que o corpo em questão é o corpo de um organismo vivo. Até mesmo a continuidade de corpo e cérebro vivos ainda

não pode ser suficiente para a preservação da identidade de Shaun. Isso é demonstrado pela possibilidade de uma amnésia total que discutimos antes (seção "Teorias reducionistas da identidade pessoal"). Ainda que um ser humano que sofreu disso tivesse um e o mesmo corpo vivo, parece que a descontinuidade mental radical excluiria a identidade da pessoa antes e depois da amnésia.<sup>2</sup>

A forma como Locke ajustou essa possibilidade foi traçando uma distinção entre um homem – isto é, um ser humano vivo – e uma pessoa. Aparentemente, após a amnésia total, teríamos um e o mesmo homem, isso é, um e o mesmo organismo humano vivo. Mas visto internamente, teríamos de dizer que o que ocorreu foi que uma pessoa deixou de existir devido à perda total da memória, e que outra pessoa iniciou na medida em que novas memórias começaram a ser armazenadas. Em outras palavras, temos duas pessoas habitando um e o mesmo corpo sucessivamente, uma após a outra.

Tendo traçado a distinção entre o conceito de um ser humano e o conceito de uma pessoa, Locke também estava consciente de outras possibilidades nas quais seria necessário que os dois conceitos se separassem. Portanto, seria possível imaginar a situação em que duas pessoas habitam um e o mesmo corpo humano, mais ou menos consecutivamente. Locke imagina um caso em que havia uma pessoa diurna que poderia lembrar o que ocorreu nos dias passados, mas não nas noites passadas; e uma pessoa noturna que, de modo inverso, poderia lembrar de eventos que ocorreram à noite, mas não durante o dia. Isso, parece, implica duas correntes de consciência incomunicáveis, manifestando-se por meio de um e o mesmo corpo. Mas se isto é uma possibilidade, então também é possível que um fluxo de consciência pudesse manifestar-se por

meio de dois corpos humanos. Finalmente, apesar de não pensarmos nos papagaios como pessoas, não é possível que pudesse existir um papagaio muito racional, assim como os cavalos racionais nas *Viagens de Gulliver*? Se for assim, não seríamos forçados a conceder que papagaios e cavalos poderiam também contar como pessoas?

A razão pela qual a distinção entre o conceito de um ser humano e o conceito de uma pessoa possa nos surpreender como estranha é a que segue. Primeiramente, as únicas pessoas que atualmente conhecemos são os seres humanos. Em segundo lugar, como casos de personalidade múltipla são distantes e raros entre nós, o volume de nossas experiências confirma o fato de que, se temos um ser humano vivo adulto e normal à nossa frente, então temos apenas uma pessoa à nossa frente. Em outras palavras, as entidades que se enquadram nos diferentes conceitos de ser humano e de pessoa, respectivamente, normalmente coincidem. Sejamos claros a este respeito. O conceito de uma pessoa não é equivalente ao conceito de um ser humano. Se fossem equivalentes, então falar sobre pessoas não-humanas seria uma contradição em termos. Mas a mera capacidade de imaginar formas de vida que não são humanas, ainda que qualificadas como pessoas, tais como nos personagens muito variados na série de ficção científica *Jornada nas Estrelas*, ajuda a estabelecer a não-equivalência dos conceitos em questão. Similarmente, é possível conceber seres humanos que, infelizmente, deixaram de

ser pessoas, como as vítimas de acidentes em coma permanente ligadas em máquinas de preservação da vida.

Antes eu disse que a continuidade física poderia vir a ser necessária para a identidade pessoal. Mas o que exatamente significa continuidade física? É natural interpretar isso significando que uma pessoa agora é uma e a mesma que uma pessoa do passado se, e somente se, essa pessoa tem um e o mesmo corpo agora que teve outrora. Dizendo isso de outra maneira, o que está por trás do meu julgamento sobre a identidade de meu estudante, Shaun, que forneceu o exemplo com o qual abri este capítulo, é a suposição plenamente natural de que a pessoa à minha frente agora, que chamo de Shaun, é fisicamente contínua com a pessoa que chamei de Shaun no passado.

Se estivesse com qualquer dúvida de que a pessoa à minha frente agora fosse uma impostora, um mero sósia, poderia em teoria traçar a trajetória do indivíduo à minha frente de volta ao longo de tempo e espaço para ver se ela levava à pessoa que chamei de Shaun na semana passada. Se assim fosse, poderíamos concluir que o Shaun à minha frente agora ainda é uma e a mesma pessoa a quem ensinei na semana passada. Se, por outro lado, a trajetória da pessoa à minha frente agora não levou ao Shaun que ensinei na semana passada, poderíamos concluir que eu tenha uma réplica do Shaun original à minha frente agora. Isso é representado diagramaticamente na Figura 10.1.

Shaun na semana passada — Trajetória contínua de um ser humano físico — Shaun agora  
retornando ao longo de tempo e espaço

Uma e a mesma pessoa

**Figura 10.1**

Porém, uma objeção radical a essa teoria foi examinada por Locke em seu experimento mental a respeito do princípio e do sapateiro. Se, diz Locke, a alma de um príncipe, carregando todas as memórias da vida do príncipe e seus traços de caráter e habilidades, entrasse no corpo de um sapateiro, e a alma do sapateiro entrasse no corpo do príncipe, todos concordariam que a pessoa a quem eles antes pensavam como sendo o príncipe era agora na realidade o sapateiro, e vice-versa.

Uma variante materialista atualizada dessa teoria foi fornecida por Sydney Shoemaker.<sup>3</sup> Estamos pensando em duas pessoas, Brown e Robinson, que estão se submetendo a operações cerebrais que envolvem a remoção de seus cérebros dos seus crânios, sendo, depois, recolocados. Infelizmente, o cérebro de Brown é equivocadamente inserido na cabeça de Robinson, e o cérebro de Robinson é colocado na cabeça de Brown. Este último paciente morre, deixando a pessoa com o corpo de Robinson contendo o cérebro de Brown, como mostrado na Figura 10.2.

Segundo Shoemaker, vamos chamar o indivíduo resultante de Brownson. Brownson, parece razoável supor, receberá um grande choque após ter-se recuperado. Ao olhar-se no espelho, verá não o que esperava ver, a saber, o rosto de Brown olhando de volta para ele, mas em seu lugar o rosto de Robinson. Parece provável que ele descobrirá que tem o corpo de Robinson, e não seu próprio corpo

original. Quando Brownson é perguntado sobre onde ele nasceu, qual seu trabalho e quem são sua esposa e seus amigos, tudo que ele disser, supomos, irá se adequar à vida de Brown, e não à de Robinson. Além disso, sua personalidade e suas habilidades também corresponderão às de Brown, e não às de Robinson. Em tal situação seria difícil evitar a conclusão de que Brownson é Brown e não Robinson. Brown teve uma troca de corpo ou, para dizer de outra maneira, virtualmente todo um transplante de corpo, com exceção do cérebro de Robinson.

Pode-se objetar que a maneira de apresentar a questão é meramente uma petição de princípios quanto à identificação de Brownson com Brown. Porém, não apenas isso está em acordo com a intuição de muitas pessoas, mas, o mais importante, é vigorosamente sustentado pelas teorias materialistas da mente. Se a mente é apenas o cérebro, como os teóricos da identidade mente/cérebro afirmam vigorosamente, ou se, com menor firmeza, a mente é mantida pelo cérebro e superveniente a ele, seja como uma coleção de propriedades não-físicas ou como uma função ou programa encarnado no hardware do cérebro, então certamente somos obrigados a concluir que Brownson é Brown, e não Robinson. De acordo com essas teorias materialistas, onde meu cérebro vai, vou eu também, seja porque sou apenas meu cérebro ou porque minha existência como um ser consciente depende de meu cérebro. Em um caso desses, a identidade do corpo de uma pessoa não é necessária nem suficiente para a identidade da pessoa. Não é necessária, pois Brown sobrevive no corpo de outra pessoa – nesse caso, o de Robinson. Temos aqui a situação em que há um corpo diferente, isto é, o de Robinson, mas uma e a mesma pessoa, isto é, Brown. Não é suficiente porque Robinson, depois que

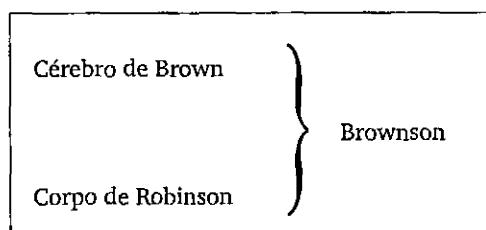


Figura 10.2

morre, sobrevive brevemente no corpo de Brown. A situação para comparação aqui seria aquela em que teríamos um e o mesmo corpo, isto é, o de Brown, e uma pessoa numericamente diferente, isto é, Robinson.

Concedamos que esse experimento mental logre destruir a continuidade da teoria corporal da identidade pessoal. Porém, não elimina a possibilidade de que a identidade da pessoa dependa da continuidade física, a saber, a continuidade física de alguma parte crucial do corpo de alguém – isto é, o cérebro da pessoa, ou até mesmo apenas algum conjunto de combinações que funcionem de modo equivalente ao cérebro. Retornarei a esse tema posteriormente.

### Fissão

Não muito depois de Shoemaker ter proposto o experimento mental do transplante de cérebro, outros filósofos, incluindo David Wiggins e Derek Parfit, viriam com uma sugestão ainda mais bizarra. Consideraram o que devemos dizer se o cérebro de uma pessoa fosse dividido em dois pela separação das fibras comissurais que mantêm os dois hemisférios do cérebro juntos, sendo então um deles colocado em um crânio apropriadamente esvaziado e o outro inserido em outro crânio vazio. Agora examine a questão apresentada no Exercício 10.4 à luz dessa sugestão.

### EXERCÍCIO 10.4

A pessoa original sobrevive à divisão ou não? Se ela sobrevive, sobrevive no hemisfério esquerdo ou direito? Ou em ambos? Qual você pensa ser a resposta correta a essa questão? Pense sobre o assunto e discuta com alguém mais antes de prosseguir a leitura.

Essa fantasia não foi trazida apenas como diversão; há um ponto sério por trás dela. Até o momento em que o experimento mental de separar o cérebro e transplantar cada uma das metades foi previsto, a principal preocupação dos filósofos, como reflete esta exposição, era com a identidade. O que significa dizer que uma determinada pessoa, P1, é numericamente idêntica (ou não, como pode ser o caso), a outra pessoa aparentemente diferente, P2? Essa questão foi particularmente premente em relação à questão da vida após a morte. Se eu morro, e não continuo a existir como uma alma fantasmagórica cartesiana, mas então Deus ressuscita meu corpo, reunindo os pedaços do que eu era feito e colocando-os de volta juntos novamente, exatamente como eles estavam quando eu estava vivo, continuo a existir ou não? A preocupação com a identidade significa que há apenas duas respostas possíveis a essa questão. Ou sou um e o mesmo que a pessoa ressuscitada, ou não sou. A identidade é uma relação de tudo ou nada: ou prevalece ou não. Não há solução pela metade ou questão de grau, tal que fizesse sentido supor que algo pudesse ser mais ou menos numericamente idêntico com alguma outra coisa. Para repetir: a respeito da identidade, ou X é numericamente idêntico a Y, ou não é.

O problema com a teoria da resurreição, na qual, na morte, o corpo é ou lentamente dissipado no túmulo ou disperso no vasto sopro pelas chamas do crematório, é que, sustentando as considerações recém-desenvolvidas na seção “A teoria da continuidade física”, a morte significa o derradeiro fim para nós. Mesmo se todas as moléculas da pessoa *pre-mortem* constituíssem o corpo do ressuscitado *post-mortem* até a medida de ocupar os mesmos lugares e papéis causais que tinham antes do falecimento da pessoa, a pessoa ressuscitada não seria uma e a mesma

que a pessoa antes da morte, mas apenas uma réplica brilhante. Dizer que isso não seria muito satisfatório seria minimizar o caso. O grande apelo das doutrinas da vida após a morte é certamente que elas supostamente dão condições para sua própria sobrevivência pessoal. Dizer que a matéria da qual você era composto voltará para o vazio, derradeiramente, para compor partes de outras pessoas, animais e plantas, não é de forma alguma confortador, precisamente porque essas outras coisas não são eu. Similarmente, ainda que eu pudesse me confortar agora que as pessoas não me esquecerão, que vou "viver" em suas memórias ou corações, ou continuarei por meio de meus filhos ou dos livros que escrevi, por fim tenho de reconhecer que minha vida chegará a um fim irrevogável e permanente, precisamente porque não sou meus filhos, meus livros ou as memórias que outras pessoas têm de mim.

Doutrinas mais místicas, que sustentam que a alma, como uma gota de água, emergirá com um grande oceano de espiritualidade e, dissolvendo-se nele, perde toda identidade e individualidade, são similarmente insatisfatórias, porque minha sobrevivência como indivíduo com uma consciência de si mesmo e de sua vida passada e presente é perdida de acordo com uma doutrina dessas.

Porém, o experimento mental do cérebro separado suscitou para considera-

ção a possibilidade de algo que não foi contemplado anteriormente: pode haver sobrevivência sem identidade. Para examinar como se supõe ser isso possível, vamos imaginar que uma das minhas atuais estudantes, Florence, é a pessoa cujo cérebro será dividido. A operação é executada e posteriormente o hemisfério esquerdo de Florence vai para um corpo à esquerda para compor a pessoa Florence(1), enquanto o hemisfério direito vai para um corpo à direita para compor a pessoa Florence(2). Após a operação ter sido executada com sucesso, tanto Florence(1) quanto Florence(2) acordam. Quando se lhes pergunta quem são, ambas dizem que são Florence. Além disso, tanto Florence(1) quanto Florence(2) fazem afirmações idênticas sobre a memória, que se ajustam exatamente à vida de Florence. Elas são, além disso, afirmações "de dentro" sobre o que Florence fez e passou – significa dizer que elas são feitas não do ponto de vista de uma terceira pessoa, observadora externa de Florence, mas, tanto quanto podemos determinar, da perspectiva privada da primeira pessoa de Florence, isto é, como a pessoa que fez ou foi submetida às coisas em questão. A situação de Florence é ilustrada na Figura 10.3.

Note que eu disse que as afirmações sobre a memória de Florence(1) e Florence(2) se ajustam à vida de Florence, e não que tanto Florence(1) quanto Florence(2)

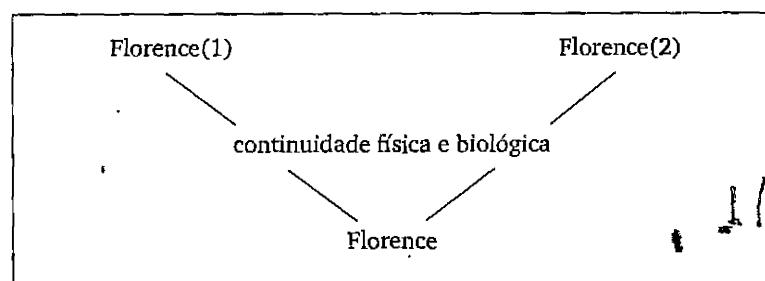


Figura 10.3 Cérebro dividido.

lemboram o que Florence fez. A razão pela qual disse isso pode ser explicada como se segue. Vamos supor que tanto Florence(1) quanto Florence(2) afirmassem lembrar de atividades anteriores a Florence ter seu cérebro dividido, que foram ensinadas por mim na semana passada. Florence(1) irá mais naturalmente expressar esse fato da memória, dizendo "Eu lembro de ter aprendido de Keith Maslin na última semana", e Florence(2) dirá o mesmo. Mas essa frase pode ser ampliada para afirmar: "Eu me lembro que Keith Maslin me ensinou na semana passada", e isto significa que a primeira ocorrência de "eu" nesta frase deve referir-se a uma e a mesma pessoa que a segunda ocorrência de "eu". Expressando de maneira mais simples o que estou dizendo, o que a frase está afirmando é que eu, Florence(1), sou uma e a mesma pessoa que Florence a quem Keith Maslin ensinou na semana passada. Semelhantemente, a frase emitida da boca de Florence(2) também afirma que eu, Florence(2), sou uma e a mesma pessoa que a Florence quem Keith Maslin ensinou na semana passada. Florence(1) tem tanto direito de fazer essa afirmação quanto Florence(2) (e vice-versa), mas o problema deve agora estar evidente. Se Florence(1) realmente é uma e a mesma pessoa que Florence, e Florence(2) de fato é uma e a mesma pessoa que Florence, então devemos ser forçados a adotar a contradição de que Florence(1) é uma e a mesma pessoa que Florence(2). Muito claramente, isso é manifestamente não verdadeiro, pois Florence(1) e Florence(2) existem em dois lugares diferentes, uma na esquerda e a outra na direita.

O princípio que leva a essa conclusão é chamado a transitividade da identidade: se X é numericamente idêntico a Y, e Y é numericamente idêntico a Z, então logicamente X deve ser numericamente idêntico a Z. Exibindo isso formalmente:

Tenha " $=$ " o significado de "é numericamente idêntico a"

Então, se

1 Florence = Florence(1), e

2 Florence = Florence(2), então

3 Florence(1) = Florence(2)

e isso é manifestamente um resultado absurdo.

O problema surge porque no uso de inglês padrão a afirmação "Eu lembro que Keith Maslin me ensinou" implica que a pessoa ou as pessoas fazendo a afirmação de memória, neste caso Florence(1) e Florence(2), são uma e a mesma pessoa que foram submetidas à coisa que é lembrada. Mas Florence(1) e Florence(2) não podem, ambas, estar lembrando, no sentido usual, no caso em questão, porque se é admitido que ambas possam legitimamente dizer que se lembram de que Keith Maslin as ensinou antes da divisão, isso levará, como vimos, por meio da transitividade da identidade, à afirmação sem sentido de que Florence(1) e Florence(2) são uma e a mesma pessoa.

Para contornar essa dificuldade, filósofos possuem um artifício simples. Em vez de dizer que tanto Florence(1) quanto Florence(2) lembram que Keith Maslin as ensinou, podemos dizer, que elas quase-lembram ou que é como se lembrassem. A quase-memória é exatamente como a memória comum de eventos, não elaborada por uma perspectiva de terceira pessoa, mas do ponto de vista de uma primeira pessoa, "de dentro", exceto que o efeito do termo "quase" é anular a implicação de que a pessoa, tendo a quase-lembrança, é uma e a mesma pessoa que executou a ação quase-lembrada. O lembrar comum por uma pessoa de algo que ela fez pode então ser visto como uma forma de quase-lembrança; efetivamente isso será um quase-lembrar, em que a pessoa tendo a quase-lembrança é uma e a mesma que a pessoa que executou a ação lembrada.

Porém, para retornar à nossa questão original, o que ocorre à Florence original quando a fissão tem lugar? Algumas respostas parecem ter sido descartadas desde o início. Primeiramente, não temos fundamento para dizer que Florence é idêntica a Florence(1) em vez de Florence(2), tampouco que Florence seja idêntica a Florence(2) em vez de Florence(1), pois ambas, Florence(1) e Florence(2), têm uma pretensão igualmente boa ao título. Em segundo lugar, somos proibidos, sob pena de contradição, de dizer que Florence é numericamente idêntica a Florence(1) e que Florence também é numericamente idêntica a Florence(2). Em terceiro lugar, parece errado dizer que Florence não mais existe. Se os dois transplantes têm sucesso, e as duas pessoas despertam com quase-memórias da vida de Florence, parece haver um sentido de que Florence ainda está por aqui. Como Parfit uma vez expressou, como pode um duplo sucesso ser equivalente a um fracasso?

Uma maneira de abordar esse problema é imaginar que você é Florence e está prestes a sofrer uma operação de divisão. Você julgaria as chances de sucesso como equivalentes à morte, ou pensaria que em algum sentido você, ou ao menos uma parte significante de você, sobreviveria? Compare o que ocorre quando o bulbo de uma tulipa se divide. A tulipa em algum sentido sobrevive, como duas metades. A divisão de um bulbo de tulipa não é equivalente a triturar o bulbo da tulipa até torná-lo uma pasta. O bulbo certamente não sobreviveria à Trituração, enquanto no processo de formar duas novas plantas há um sentido no qual ele efetivamente sobrevive.

Da mesma maneira, somos convidados a pensar que a fissão não significa o fim de uma pessoa. Florence sobrevive tanto como Florence(1) quanto como Florence(2), mesmo que, sob pena de contradição, não se pode afirmar que Florence seja

numericamente idêntica a ambas. Como Florence(1), Florence sobrevive com um depósito de quase-memórias de sua intacta vida passada. Do ponto de vista de Florence(1), as quase-memórias não são diferentes das memórias genuínas da vida de Florence, pois elas foram produzidas essencialmente da mesma maneira que as memórias comuns são produzidas quando nenhuma fissão está envolvida. Similarmente, Florence também sobrevive como Florence(2), que, de seu ponto de vista, também possui um estoque de quase-memórias que se ajustam à vida de Florence e são tão boas quanto as reais. (Em um sentido, a partir das perspectivas de primeira pessoa de Florence(1) e Florence(2), as quase-memórias da vida de Florence são reais, porque são fenomenologicamente indistinguíveis das memórias genuínas.)

O que o caso da fissão parece estabelecer é que podemos ter sobrevivência pessoal sem identidade. Não importa que não possamos ter identidade. O que realmente nos interessa é a sobrevivência, e o exemplo da divisão é conservador da pessoa, pois nos dá o que queremos, isto é, a continuidade de alguma forma, mesmo que nenhum dos produtos da divisão, Florence(1) e Florence(2), possa afirmar ser es- trita e numericamente idêntico a Florence.

Um aspecto importante do caso da fissão apenas esboçado é que temos tanto continuidade física quanto psicológica. Realmente, pareceria nesse caso que há continuidade psicológica apenas porque temos continuidade física. Os processos físicos que ocorrem em cada metade do cérebro mantêm a vida mental de cada uma das pessoas resultantes, Florence(1) e Florence(2), assim como os processos em todo o cérebro não dividido sustentam a vida mental da pessoa original, Florence.

Isso sugere a questão: podemos imaginar uma situação em que temos continuidade psicológica sem continuidade

física do cérebro? Se sim, então talvez possamos sobreviver à morte e destruição de nossos cérebros. Isso nos leva, por uma progressão natural, à continuidade psicológica como uma teoria da identidade pessoal, o que compõe o tema do próximo capítulo.

### **Questões para reflexão**

1. Posso me lembrar de ter sido alguém diferente de quem sou agora?
2. Que dificuldades filosóficas estão envolvidas na suposição de que, após minha morte, Deus me ressuscitará?
3. A possibilidade de uma fissão cerebral estabelece que o que importa para a sobrevivência pessoal é a continuidade física, e não a identidade?

### **Sugestões para leituras profundadas**

- Não parece existir sobre a identidade pessoal muitos textos escritos como introduções para o tópico.

CARRUTHERS, P. *Introducing Persons: Theories and Arguments in the Philosophy of Mind*. London and Sydney, Croom Helm, 1986 ajuda a preencher essa lacuna, e as partes relevantes do livro são recomendadas para os iniciantes nessa área da filosofia.

- SHOEMAKER, S. *Self-Knowledge and Self-Identity*. Cornell, Cornell University Press, 1963 é o livro que alavancou uma renovação do interesse na identidade pessoal atualmente. Ele é altamente recomendável.
- WILLIAMS, B. *Problems of the Self*. Cambridge, Cambridge University Press, 1973 apresenta contribuições importantes e originais ao debate, reunindo muitos de seus importantes artigos. É uma leitura essencial para todos os interessados na identidade pessoal.
- SHOEMAKER, S. e SWINBURNE, R. *Personal Identity*. Oxford, Blackwell, 1984 traz um debate interessante e vivo entre dois pensadores influentes, defendendo diferentes abordagens à identidade pessoal. É altamente recomendável.

## NOTAS

---

1. WILLIAMS, B. "Personal Identity and Individuation", reimpresso em WILLIAMS B., ed. *Problems of the Self, Philosophical Papers, 1956-72*. Cambridge, Cambridge University Press, 1973, p. 1-19.

2. Alguns filósofos discordariam. Por exemplo, Leibniz, REMNANT, P. e BENNETT, J., eds. e trads. *New Essays concerning Human Understanding*, Cambridge, Cambridge University Press, 1992, parágrafos 236-7, escreveu o seguinte sobre essa possibilidade:

Portanto, se uma doença interrompeu a continuidade de meu vínculo de consciência de tal forma que não sei como cheguei ao meu estado presente, mesmo que possa lembrar coisas passadas, o testemunho de outros pode preencher a lacuna em minhas recordações. Posso até mesmo ser punido sob esse testemunho se fiz algum mal deliberado durante um intervalo que essa doença me fez esquecer pouco tempo depois. E se eu esquecera todo o meu passado, e precissei que me ensinassem tudo de novo, até mesmo meu nome e como ler e escrever, ainda poderia aprender com os outros sobre minha vida durante o estado precedente; e, similar-

mente, teria mantido meus direitos sem ter de ser dividido em pessoas e ser feito herdeiro de mim próprio.

Compare com MACKIE, J. L. *Problems from Locke*. Oxford, Clarendon Press, 1976, p. 190:

Memória e caráter não são alternativas de evidência para a continuidade corporal; servem, no direito e em todos os casos ordinários, como evidência da continuidade pessoal somente na medida em que são evidência da continuidade corporal. Tal continuidade ocupa uma posição similar como evidência até da minha própria identidade como pessoa. Sempre e em todo lugar no passado, o mesmo homem que, como eu, era consciente e ativo, presumo que era eu quem estava lá e que fez o que quer que seja que este homem tenha feito. Estou bastante pronto a atribuir a mim mesmo inumeráveis ações e experiências que não recordo agora, uma vez que esteja satisfeito da continuidade corporal dessa pessoa comigo.

3. SHOEMAKER, S. *Self-Knowledge and Self-Identity*. Cornell, Cornell University Press, 1963, c. 1, § 8-10.

# 11

## IDENTIDADE PESSOAL COMO CONTINUIDADE PSICOLÓGICA

### Objetivos

Como resultado da leitura deste capítulo você deve:

- Entender a abordagem de Locke da identidade pessoal como memória e as objeções de Reid e Butler que podem ser levantadas contra ela.
- Entender a abordagem do encadeamento psicológico e da continuidade de Parfit sobre a identidade pessoal e as objeções que podem ser levantadas contra ela.
- Ser capaz de avaliar criticamente a teoria do feixe do eu de Hume.
- Compreender como a sobrevivência sem identidade é supostamente possível e ser capaz de avaliar criticamente essa afirmação.
- Entender como a continuidade psicológica sem continuidade física é supostamente possível e ser capaz de avaliar criticamente essa afirmação.
- Entender a afirmação de Unger de que a continuidade física na forma da existência continuada do cérebro é essencial para a preservação da identidade pessoal e ser capaz de avaliar criticamente essa afirmação.
- Entender as diferenças entre as abordagens reducionista e não-reducionista da identidade pessoal e ser capaz de avaliar criticamente essas opções.

### Introdução

John Locke (1632-1704) foi o criador da teoria de que a identidade pessoal consiste na continuidade psicológica. Para fornecer um resumo da teoria, nada melhor do que citar a própria afirmação de Locke:

Pois, sendo isso a mesma consciência que faz um homem ser ele mesmo para si mesmo, a identidade pessoal depende disso apenas, quer seja anexado apenas à sua substância individual, quer possa ser continuada em uma sucessão de muitas substâncias. Pois, na medida em que qualquer ser inteligente pode repetir a ideia de qualquer ação passada com a mesma consciência que teve dela inicialmente,

e com a mesma consciência que tem de qualquer ação presente, desde que seja o mesmo eu pessoal. Pois é por meio da consciência que ele tem de suas ações presentes que ele é eu para si próprio agora, e então será o mesmo eu na medida em que a mesma consciência possa expandir-se para ações passadas ou futuras. Não seria através de um distanciamento temporal, duas pessoas assim como um homem não seria duas pessoas por vestir hoje diferentes roupas daquelas que vestiu ontem, com um longo ou breve sono entre elas: *a mesma consciência unindo aquelas ações distantes dentro da mesma pessoa, quaisquer que sejam as substâncias que contribuíram à sua produção.*<sup>1</sup>

A essência da abordagem lockeana é essa: a identidade pessoal não consiste nem na persistência de uma porção organizada de matéria – o corpo humano ou uma parte importante deste corpo, o cérebro – nem na continuidade de uma alma cartesiana imaterial, mas, em uma certa espécie de continuidade e encadeamento entre uma série de experiências. As substâncias nas quais essas experiências são encarnadas são irrelevantes para a identidade da pessoa. Essa é uma reminescência do funcionalismo. O que faz com que estados mentais sejam o que são não é a coisa, seja material ou até mesmo, é possível, imaterial (veja a "Introdução" no Capítulo 5), que eventualmente os incorpora, mas as relações de *input/output* que só podem ser especificadas formalmente. Similarmente, o que Locke está afirmando é que, desde que haja uma série encadeada de experiências, temos identidade pessoal, mesmo se essa cadeia de estados mentais fosse transferida de uma substância física, ou mesmo de uma substância imaterial, para outra.

Dizendo de modo mais preciso, Locke está afirmando que uma pessoa no presente (tempo  $t_1$ ) é uma e a mesma que uma pessoa no passado (tempo  $t_0$ ), e, e somen-

te se, a pessoa no presente pode lembrar o que a pessoa no passado fez: "Pois, na medida em que qualquer ser inteligente pode reproduzir a ideia de uma ação passada com a mesma consciência que dela teve inicialmente, e com a mesma consciência que tem de qualquer ação presente, desde que seja o mesmo eu pessoal".<sup>2</sup>

Porém, isso suscita imediatamente a seguinte questão: a memória de um evento passado, digamos uma ação que você praticou, é necessária para fazer com que a pessoa que está recordando seja uma e a mesma pessoa que fez a ação? Pense sobre isso, e examine o Exercício 11.1, antes de prosseguir a leitura.

### EXERCÍCIO 11.1

Uma pessoa X é uma e a mesma que uma pessoa Y apenas se X pode lembrar, de uma perspectiva de primeira pessoa, o que Y fez?

### A objeção de Reid ao argumento de Locke

A primeira objeção que pode lhe ocorrer é esta: certamente você pode esquecer que fez certas coisas, e, no entanto, isso não significa que você não seja uma e a mesma pessoa que as fez. Posso ter esquecido que fui eu, e não minha esposa, quem colocou o lixo na rua na última semana, e ainda assim, não obstante isso, sou uma e a mesma pessoa que colocou o lixo na rua. Mais dramaticamente, Ivan Demanyuk pode genuinamente insistir que ele não tem lembrança de haver cometido horrores nos campos nazistas de extermínio, e apesar de tudo isso ele pode ter sido a pessoa que perpetrhou aquelas atrocidades.

Essa objeção foi desenvolvida por Thomas Reid (1710-1796) por meio da se-

guinte história. Considere um soldado que uma vez foi um menino que roubou maçãs de um pomar e, consequentemente, foi chicoteado. Posteriormente, porém, esse mesmo menino tornou-se um bravo oficial que capturou um soldado em batalha.

Posteriormente, o oficial foi promovido e, por fim, tornou-se um famoso general. Se a memória supostamente é necessária para a identidade pessoal, temos o seguinte resultado:

- O bravo oficial pode lembrar de ter roubado maçãs, então ele é uma e a mesma pessoa que o menino que as roubou e foi chicoteado.
- O general pode lembrar de ter aprisionado o soldado, então ele é uma e a mesma pessoa que o bravo oficial.
- O general, porém, não pode lembrar-se de ter roubado maçãs quando menino, então ele não é uma e a mesma pessoa que o menino.

Agora examine o Exercício 11.2.

### **EXERCÍCIO 11.2**

O que você diria que o contra-argumento de Reid mostra? Reflita sobre sua resposta antes de prosseguir a leitura.

Pareceria que se o general é o bravo oficial (b) e que o bravo oficial é o menino (a), então o general é o menino (através da transitividade da identidade: se  $x = y$ , e  $y = z$ , então  $x = z$ ). Mas, de acordo com (c), o general não é o menino, porque ele não pode lembrar o que o menino fez. Consequentemente, o general tanto é, como não é, uma e a mesma pessoa que o menino. Mas isso é uma contradição evidente, consequentemente precisamos rejeitar a memória como condição necessária da identidade pessoal.

### **Lidando com a objeção de Reid**

Locke estava no caminho certo, de acordo com o filósofo contemporâneo Derek Parfit, mas exagerou seu caso ao insistir que uma pessoa no presente, X, não poderia ser idêntica a uma pessoa no passado, Y, que cometeu algum crime, a menos que X se lembre de haver cometido o crime. Parfit sugere que X pode ser Y, contanto que haja o suficiente do que ele chama de continuidade e encadeamento – C & E – entre as memórias, assim como entre outros aspectos psicológicos. Por encadeamento Parfit quer dizer conexões diretas da memória: se X pode lembrar o que Y fez há muitos anos atrás, a partir de seu íntimo, então X é idêntico a Y. Mas mesmo se conexões diretas estão faltando, ainda pode haver continuidade. A ideia aqui é de que há uma cadeia de memórias diretas sobrepostas. Então, no domingo posso lembrar algo do que fiz no sábado, mas talvez nada do que fiz na sexta. Porém, no sábado posso lembrar algo do que fiz sexta, apesar de não poder lembrar nada do que fiz na quinta. Na sexta, entretanto, posso lembrar algo que fiz na quinta, mas nada do que fiz na quarta... e assim por diante. Revisando o critério de Locke dessa maneira, podemos dizer que X é idêntico a Y porque há suficiente continuidade e encadeamento entre eles. Podemos ver a proposta representada diagramaticamente na Figura 11.1.

O corolário dessa abordagem é que, se não há de todo C & E de aspectos psicológicos, como no caso da amnésia total imaginado no Capítulo 10, seção “Teorias reducionistas da identidade pessoal”, então parece que estamos autorizados a concluir que uma pessoa no presente não é uma e a mesma que uma pessoa no passado, mesmo que haja continuidade física em termos de persistência de corpo e cérebro.

A modificação que Parfit faz de Locke pode acomodar outras espécies de C & E psicológicos além da memória, por exemplo, a conexão entre uma intenção de executar uma ação específica e a posterior execução deste ato. Também há conexões entre crenças e desejos passados e estados da mente no presente, tanto como conexões entre traços de personalidade, hábitos e habilidades no passado e no presente.

Parfit salienta que, como a identidade é uma relação transitiva (se  $x = y$  e  $y = z$ , então  $x = z$ ) e o encadeamento não é uma relação transitiva (se na segunda lembro o que fiz na terça, e na terça lembro o que fiz na quarta, disto não se segue que na segunda lembro o que fiz na quarta), então o encadeamento sozinho não pode contar como um critério de identidade. O que é também necessário é a continuidade.

Além disso, Parfit faz a distinção entre encadeamento forte e fraco, porque encadeamento é uma questão de grau. Para X ser idêntico a Y deve haver conexões psicológicas suficientes entre X e Y, em outras palavras, precisa haver aí um encadeamento forte. Embora "suficiente" não possa ser precisamente definido, Parfit sustenta que pode ser dito que X é numericamente idêntico a Y se o número

de conexões em qualquer dia é ao menos a metade da quantidade de conexões realizadas ao longo de todo o dia nas vidas de todas as pessoas reais.

Agora Parfit está apto a usar a noção de encadeamento forte na definição de identidade pessoal. Diferente do encadeamento, a continuidade é uma relação transitiva. Se X na segunda é contínuo com Y na terça, e Y na terça é contínuo com Z na quarta, então X é contínuo com Z. Parfit sugere a seguinte definição de identidade pessoal que ele chama de critério psicológico:

- (1) Há continuidade psicológica se, e somente se, houver cadeias sobrepostas de encadeamento forte. X hoje é uma e a mesma pessoa que Y em algum tempo passado se e somente se (2) X é psicologicamente contínuo com Y, (3) esta continuidade tem o tipo certo de causa, e (4) não adotou uma forma ramificada. (5) A identidade pessoal ao longo do tempo consiste simplesmente na manutenção de fatos como (2) e (4).<sup>4</sup>

Nesse estágio, porém, o significado das condições (3) e (4) e a razão para sua inclusão podem não ser muito claros. Porém, irei brevemente pospor sua explicação, pois é melhor que essa tare-

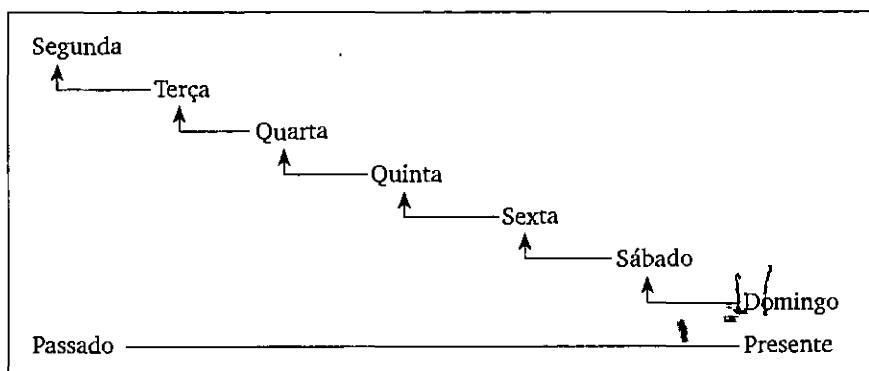


Figura 11.1 Cadeia de memórias sobrepostas.

fa aguarde até que eu tenha lidado com outras objeções ao modelo Locke/Parfit de identidade pessoal como continuidade psicológica.

### A objeção de Butler ao argumento de Locke

Certamente, se X pode genuinamente lembrar de maneira costumeira, de dentro, o que Y fez, então se segue que X é Y. Se eu agora posso lembrar de ter ensinado Florence na última semana, então devo ser um e o mesmo que a pessoa que ensinou Florence na última semana. A razão, como vimos antes, é óbvia. "Eu lembro de ter ensinado Florence na última semana" é elíptico com "Eu lembro que ensinei Florence na última semana". Para essa sentença fazer sentido, a referência das duas ocorrências do pronome "eu" deve ser a uma e à mesma pessoa o tempo todo. Mas isso leva a uma objeção muito reveladora, inicialmente formulada pelo Bispo Joseph Butler (1692-1752), a saber, de que a abordagem de Locke é circular. Butler escreveu: "Alguém deveria realmente pensar que é autoevidente que a consciência da identidade pessoal pressupõe, e portanto não pode constituir, a identidade pessoal, muito menos que o conhecimento, em qualquer outro caso, pode constituir a verdade que ele pressupõe".<sup>5</sup>

Em outras palavras, é realmente verdadeiro que lembro de ter ensinado Florence na última semana, então isso pressupõe que sou uma e a mesma pessoa que ensinou Florence. Se não fosse eu quem ensinou Florence, então estou equivocado, e, portanto, não posso, afinal, realmente estar lembrando isso. Minha identidade numérica com a pessoa que ensinou Florence está sendo usada como um critério sobre se tenho ou não uma memória ge-

núina de tê-la ensinado. É precisamente dessa maneira que a noção de memória genuína pressupõe a identidade pessoal e por que a memória não pode ser invocada a fim de constituir a identidade pessoal.

### Lidando com a objeção da circularidade

A elaboração de Parfit da teoria de Locke ainda permite a objeção de circularidade. Para evitar esse problema, Parfit sugere que a noção de memória seja substituída pela de quase-memória, como já foi definida (Capítulo 10, seção "Fissão"). Tenho uma quase-memória de ensinar Florence na última semana, uma memória aparente de instruí-la, mas isso deixa em aberto se realmente era eu que a ensinava ou alguma outra pessoa. Portanto, a identidade da pessoa realizando o quase-lembrar não é pressuposta, e a circularidade na análise é contornada.

Para tornar a análise modificada de Parfit tão clara quanto possível, examine o seguinte caso da vida real. Lembro que aos 3 anos meu pai e eu presenciamos o lançamento de um barco salva-vidas em Exmouth, Devon. Qualquer tentativa de análise de minha identidade ao longo dos anos que decorreram desde que tinha 3 anos que use a noção de memória pressupõe a identidade pessoal, como vimos; consequentemente, façamos a análise, por outro lado, em termos de quase-memória. Assim,

1. Parece que me lembro do lançamento do barco salva-vidas.
2. Alguém viu o lançamento do barco salva-vidas.
3. Minha memória aparente é dependente, da maneira correta, dessa experiência passada.

Esta fórmula deixa em aberto que o alguém em (2) poderia não ser eu mesmo.

Poderia quase-lembrar as experiências de outra pessoa. Provavelmente será objetando que isso nunca acontece. De fato, eu efetivamente nunca me lembro das experiência de alguma outra pessoa "a partir de dentro", isto é, de uma perspectiva da primeira pessoa, em vez de uma da terceira pessoa. Isso é verdade, mas deve necessariamente ser assim?

Suponha que a memória é de alguma maneira codificada nas células cerebrais de uma pessoa, e cirurgiões desenvolvem uma técnica para transplantar essas células de uma pessoa para outra. Isso poderia capacitar ao receptor acessar as memórias do doador a partir de dentro; assim, parece ao receptor que ele estava lembrando das experiências do doador a partir de dentro. A experiência de parecer lembrar das experiências do doador poderia assemelhar-se à lembrança de suas próprias experiências, à exceção de que, de fato, seria quase-lembrar-se das experiências de outra pessoa. O fato de que elas eram experiências de outra pessoa não entraria no modo de apresentação dessas experiências – as experiências seriam apresentadas a mim como se elas fossem minhas próprias experiências originais.

Em termos de parecer que eu lembre ter visto o lançamento do barco salva-vidas, o que poderia ter ocorrido é isso. Ao longo de muitos anos, perdi totalmente todas as memórias de ver o barco ser lançado, mas meu pai não. Subsequentemente, tive algumas de suas células de memória implantadas em mim, e como resultado me parece que estou lembrando o lançamento do barco. Na realidade, estou somente quase-lembrando o lançamento do barco, uma vez que é a experiência de meu pai de estar vendo o barco saindo do galpão que está causando minha aparente memória do evento, e não minha experiência de ver o barco lançado que está causando minha memória disso.

O que esse exemplo levemente fantástico propõe suscitar é que a memória é uma noção causal. O mesmo se aplica à quase-memória. Para uma experiência presente contar como uma experiência de memória, ou mesmo uma experiência da quase-memória, aquela deve ser causada pela experiência original de alguém, seja da própria experiência no caso da memória direta comum, ou de alguém mais, no caso da quase-memória. Se o que parece ser uma experiência de memória foi causado, digamos, pela injeção de uma droga ou uma batida na cabeça, então aquilo não pode contar como se fosse uma lembrança comum, nem mesmo como quase-lembrança. A experiência não será de qualquer tipo de memória em absoluto.

Porém, para retornar ao argumento principal, lembremo-nos que o que estamos procurando é uma abordagem não-circular da identidade pessoal estruturada em termos da noção causal de quase-lembrança, de uma experiência passada causando outra experiência presente, em que essa relação é expressa de tal modo que a identidade da pessoa, cujas experiências estão em questão, não é pressuposta. As três afirmações apresentadas anteriormente podem parecer ter captado o que é necessário, mas infelizmente há um problema com a explicação detalhada do que é requerido pela condição (3).

Para compreender a dificuldade, vamos considerar uma modificação do exemplo do barco salva-vidas. Suponha que, tendo visto o barco ser lançado, meu pai esqueceu totalmente esta experiência. Eu, porém, não esqueço, e alguns anos depois lembro-o do evento. Isso tem o efeito de reavivar sua memória. Porém, algum tempo depois disso, eu mesmo esqueço totalmente a experiência original. Passaram-se muitos anos até que um dia meu pai me contou como nós vimos o barco salva-vidas ser lançado. Um ano depois do últi-

mo encontro, esqueci que ele me dissera isso. Porém, como resultado de seu relato, subsequentemente isso me pareceu como se eu realmente lembrasse de ter visto o lançamento do barco salva-vidas. Na realidade, porém, eu não lembro. Isso porque minha impressão de memória resultou de suas palavras, e não diretamente nem de minha experiência de ver o barco ser lançado, nem mesmo de sua experiência do evento. Portanto, não estou genuinamente recordando a experiência, nem mesmo quase-relembmando-a. Apenas parece para mim que estou. Contudo, minha experiência original de ver o lançamento do barco salva-vidas desempenhou um papel na produção de minha memória aparente, e consequentemente as condições (1), (2) e (3) são satisfeitas. O problema é que (3) não é satisfeito da maneira correta. O que precisávamos para (3) era que minha experiência de ver o lançamento do barco salva-vidas causasse diretamente a quase-memória dele. Na abordagem presente ela não o faz, porque a cadeia causal alcançada a partir da experiência original para minha quase-memória aparente disto toma uma rota incerta por meio de um agente externo, meu pai.

A dificuldade não pode, porém, ser evitada pela especificação de que, para uma quase-memória aparente ser considerada como uma quase-memória genuína e, portanto, para qualificar para inclusão na análise dá identidade pessoal, a cadeia causal que se estende da experiência original à sua quase-memória precisa ocorrer internamente por meio de uma e a mesma pessoa, e não passar externamente por outras pessoas. A razão pela qual essa estipulação não pode ser feita deve ser clara: que ela emprega o próprio conceito de uma e a mesma pessoa, que se espera que seja explicado de uma maneira não-circular.

Vamos considerar novamente o que Parfit está tentando fazer. O que parece emergir, afinal, é que o objetivo de Parfit é fornecer uma teoria da identidade pessoal puramente em termos de relações causais que prevaleça entre experiências e outros aspectos psicológicos. Em última análise, isso é tudo que a identidade pessoal realmente significa para Parfit, uma cadeia de experiências e outros aspectos psicológicos causalmente relacionados uns aos outros, de maneira correta, isto é, "direta", não-incerta. Porém, parece duvidoso que mesmo essa análise modificada possa evitar a importação da circularidade para dentro dela.

Porém, seja isso o que for, é de algum interesse notar que a abordagem de Parfit possui um precedente histórico, por quanto possui mais do que uma semelhança passageira com a teoria do feixe do eu, de Hume. A teoria de Hume prosseguia da seguinte maneira. Ele queria deixar claro o que o sujeito de experiências, o "eu" ou o próprio eu\* ao que todas elas pertencem, realmente existia. Esse sujeito de experiências, ao que parece, não poderia meramente ser ele mesmo outra experiência, mas deveria ser colocado sobre e além delas como um diferente tipo de item, o tipo de coisa capaz de possuir experiências, mas que não é redutível a qualquer uma delas, ou, de fato, mesmo à sua totalidade.

Hume acreditava que todos os conceitos significativos ou ideias, como ele os chamava, afinal de contas derivavam da experiência sensível que supria o que ele chamava de "impressões". Em conformidade com isso, para Hume, a maneira de esclarecer algum conceito que seja obscuro é rastreá-lo de volta às suas origens nos sentidos. A ausência de clareza sobre o eu, pensava ele, significava que alguém deve tentar encontrar a impressão

\* N. de T.: No original: *self*.

sensorial a partir da qual supostamente o conceito de eu se deriva como pálida imagem. A fim de encontrar essa impressão sensorial, Hume faz uma introspeção para encontrar o eu que é o possuidor das experiências. Porém, de acordo com seu famoso relato, não pôde encontrar isso. O máximo que pôde descobrir foram experiências particulares que ele estava tendo e que mudavam rapidamente. Ele então foi levado a concluir pela inexistência do eu. Não há sujeito de experiências, não há "eu" além das próprias experiências. O eu, ele então concluiu, deve ser, em última instância, nada mais do que um feixe de experiências relacionadas das maneiras apropriadas. Mas o que são essas maneiras? Se considerarmos seis experiências, E1, E2, E3, E4, E5 e E6, por que princípio decidiremos se temos apenas um eu consistindo de seis experiências, ou dois eus, cada um consistindo de três experiências, ou três eus, cada um consistindo de duas experiências, ou mesmo seis eus, cada um consistindo de apenas uma experiência?

Hume observou que havia apenas três maneiras pelas quais elas podem ser relacionadas umas às outras. Primeiramente, poderiam ser contíguas, isto é, próximas umas às outras. Em segundo lugar, poderiam se parecer umas às outras. Em terceiro lugar, poderiam estar causalmente relacionadas. Qual dessas relações, se é que há alguma, pode ser usada para decidir quais experiências devem constituir qual eu? Pense sobre essa questão e depois examine a pergunta do Exercício 11.3.

### **EXERCÍCIO 11.3**

Que dificuldades você encontra na proposta de Hume? Discuta isso com alguém mais antes de prosseguir a leitura.

Para começar, a contigüidade não parece ser muito útil porque não está claro em que sentido as experiências possam ser contíguas umas às outras em oposição a que as experiências são experiências. O próprio exemplo de Hume em seu *Enquiry concerning Human Understanding*\* é de apartamentos que estão próximos um do outro. Porém, enquanto os quartos podem literalmente estar contíguos um ao outro, é difícil atribuir sentido à afirmação de que as experiências das pessoas que estão vendo os quartos estão próximas umas das outras, a menos que isto seja meramente uma maneira obscura de dizer que duas pessoas de carne e ossos, cada uma das quais está tendo uma experiência de ver um dos quartos, estão situadas próximas uma da outra.

Em segundo lugar, a similaridade de experiências não é nem necessária nem suficiente para determinar quais delas pertencem a quais pessoas. Não é necessária, pois é possível que alguém possa ter um fluxo de experiências, nenhuma das quais se pareça com as outras. Não é suficiente porque as experiências de pessoas diferentes poderiam parecer-se umas com as outras.

Isso posto, vejamos a causalidade. É necessário a uma experiência ser causalmente vinculada a outras experiências para contar como parte do feixe que constitui o eu? A ideia de uma experiência particular, que não possui relação causal com qualquer outra experiência e, contudo, conta como a experiência de uma e da mesma pessoa não parece ininteligível. Mas qualquer que seja a verdade na decisão dessa difícil questão, uma relação causal entre uma experiência e outra não parece ser suficiente para fazê-las parte de um e do mesmo eu. Isso porque faz sentido supor que a experiência de uma

\* N. de T.: Em português: *Investigação sobre o entendimento humano*.

pessoa pode dar origem à de outra pessoa – minha tristeza, por exemplo, pode fazer com que você se sinta triste – e, contudo, apesar do vínculo causal entre as duas experiências, minha experiência de tristeza ainda seria minha, enquanto a sua continuaria sua. Voltando ao caso da experiência de meu pai ver o lançamento do barco salva-vidas e, após algumas de suas células cerebrais terem sido colocadas no meu cérebro, isso supostamente causar minha quase-memória do lançamento; o fato de que sua experiência original causou minha posterior quase-memória não é suficiente para classificar a experiência original como minha em vez de ser dele. De fato, como já salientei, falar de quase-lembrar as experiências de alguém mais, mas genuinamente lembrar as minhas próprias experiências, pareceria estar invocando as noções correlativas de pessoas que são numericamente as mesmas, mas são diferentes. Isso é, naturalmente, a própria noção à qual supostamente estamos fornecendo uma análise de pureza em termos de encadeamento e continuidade das experiências e outros aspectos psicológicos. A ideia da quase-memória de Parfit é desenhada para eliminar a circularidade. No entanto, afinal de contas, ela retorna. Como Colin McGinn competentemente observou, parece que as teorias causais raramente podem atingir suficiência.<sup>6</sup>

Parfit, eu penso, ficaria imperturbável com essas críticas. Parece-me que o resultado de sua posição é esse. Existem corpos humanos e cérebros, e também existem experiências. Algumas dessas experiências são causalmente relacionadas umas às outras. Quando experiências suficientes são relacionadas umas às outras por encadeamento forte e continuidade no sentido explicado anteriormente (veja seção “Lidando com a objeção de Reid”), então temos um (o) sujeito de experiências. Como o próprio Parfit diz:

Porque atribuímos pensamentos a pensadores, é verdade que pensadores existem. Mas pensadores não são entidades que existem separadamente. A existência de um pensador apenas envolve a existência de seu cérebro e corpo, o fazer de seus feitos, o pensar de seus pensamentos, e a ocorrência de certos outros eventos físicos e mentais. Poderíamos, portanto, redescriver a vida de qualquer pessoa em termos impessoais. Explicando a unidade dessa vida, não precisamos afirmar que essa é a vida de uma pessoa em particular. Poderíamos descrever que, em tempos diferentes, foi pensado e sentido e observado e feito, e como estes vários eventos estavam inter-relacionados. Pessoas seriam mencionadas aqui apenas em termos de descrições do conteúdo de vários pensamentos, desejos, memórias e assim por diante. As pessoas não precisam reivindicar serem os pensadores de qualquer um desses pensamentos.<sup>7</sup>

Poderiam existir algumas experiências que são causalmente relacionadas a esses vínculos por continuidade forte e encadeamento, mas que não exibem em si próprias continuidade forte e encadeamento. Mas essas experiências vinculadas com menos força não poderiam contar como memórias de um único eu. Em vez disso, talvez constituíssem quase-memórias, isto é, memórias, a partir de dentro, das experiências de outras pessoas. Além de minhas experiências, que são minhas experiências na visão de Hume/Parfit precisamente porque exibem continuidade forte e encadeamento, e portanto compreendem o núcleo central de mim mesmo, existirão algumas outras experiências que são causalmente vinculadas a esse núcleo, porém de maneira mais tênue. Essas experiências extras incluirão quase-memórias, tais como minha quase-memória de ver o lançamento do barco salva-vidas que se originou na experiência original de meu pai no caso em que suas células cerebrais

foram implantadas em mim. Há, inevitavelmente, certa indeterminação sobre quão fortemente vinculadas as experiências têm de ser para serem consideradas como parte da história de uma pessoa. Isso significa, como Anthony O'Hear apontou, que não há, na visão de Parfit, "nenhum critério claro para decidir quando um eu termina e outro começa".<sup>8</sup>

Isso leva Parfit ao que ele julga ser uma visão confortadora da morte. Após morrer, explica Parfit, não haverá um ser vivo que seja ele. Haverá, porém, muitas experiências, nenhuma das quais será vinculada às suas experiências presentes por tais conexões diretas como aquelas atualmente existentes no caso de suas memórias, ou que estão envolvidas quando uma intenção presente é executada por um ato posterior. Mas algumas experiências futuras serão relacionadas às suas experiências presentes de maneiras menos diretas. Algumas memórias serão de sua vida, e haverá alguns pensamentos e diretrizes influenciados pelos pensamentos de Parfit quando ele estava vivo. "Minha morte", ele diz, "romperá as relações mais diretas entre minhas experiências presentes e futuras, mas não romperá várias outras relações [...] Em vez de dizer 'Estarei morto', devo dizer 'Não haverá futuras experiências que serão relacionadas, de determinadas maneiras, às presentes'".<sup>9</sup> Essa maneira de redescrivêr a morte torna-a menos incômoda para Parfit, embora ele observe que para Hume torna-a muito pior, lançando-o na "mais deplorável condição imaginável, cercada com a mais profunda escuridão"<sup>10</sup>, para a qual a única cura seria jantar e jogar gamão com os amigos. Ao contrário, para Parfit isso remove a parede de vidro entre ele e os outros e o torna menos preocupado a respeito de sua própria morte. Diz ele que todo o seu legado resulta em que certas experiências, após dado tempo, não mais serão relacionadas a suas

experiências presentes, de certas maneiras. E que, ele pensa, isso realmente não pode ter tanta importância.

### **Identidade e sobrevivência sem continuidade física**

A teoria de Parfit o capacita a lidar com relativa facilidade com o que ocorre no caso de cérebros divididos. As experiências fornecidas pelos dois hemisférios após a fissão serão vinculadas por fortes cadeias de continuidade e encadeamento às experiências que ocorreram antes da separação. Embora, por causa da transitividade da identidade, nenhum dos dois produtos da divisão possa ser literalmente idêntico à pessoa original, os dois conjuntos de experiências que compreendem cada um dos produtos da divisão serão solidamente contínuos às experiências originais da pessoa. Portanto, há um sentido no qual a pessoa original não é aniquilada, mas continua a existir como dois fluxos de experiências causalmente relacionadas. Um rio que se divide, similarmente, não desaparece, mas continua a existir como dois cursos d'água separados e individuais que são contínuos ao rio original. Dessa maneira pode haver sobrevivência sem identidade.

Há, além disso, Parfit afirma, outra maneira pela qual a sobrevivência sem identidade é possível, mas que não envolve cérebros divididos ou até mesmo qualquer reparo com aspectos físicos da pessoa em absoluto. Interessantemente, por razões que em breve se tornarão claras, essa outra possibilidade é muito próxima à sugestão original de Locke de que a identidade pessoal/nada tem a ver com a identidade de substâncias materiais ou imateriais, mas é definível de maneira puramente formal, como Parfit sustenta, em termos de relações entre experiências.

Para avaliar o ponto que Parfit está elaborando, somos convidados a imaginar um equipamento chamado de equipamento de transferência de estado cerebral. Ele consiste em um cubículo no qual você entra e então é escaneado, as posições relativas das moléculas em seu corpo são registradas pela máquina. Seu corpo e cérebro, então, são vaporizados, após o que a informação armazenada na máquina é transmitida a um planeta distante, onde uma réplica molécula por molécula é criada a partir do material que está na superfície do planeta. Essa pessoa parece-se exatamente como você e tem todas as suas capacidades, habilidades, traços de caráter e hábitos. Além disso, tudo o que ele afirma lembrar se ajusta à sua vida em todos os detalhes. Antes de prosseguir sua leitura, examine a questão do Exercício 11.4.

#### **EXERCÍCIO 11.4**

A pessoa replicada é idêntica a você, ou meramente uma duplicata brilhante? Qual é sua reação a esse experimento mental? Considere razões favoráveis e contrárias à afirmação de que a réplica é você.

De acordo com Parfit, a pessoa é você. Afinal, tudo o que realmente importa para a sua sobrevivência está no lugar, a saber, continuidade e encadeamento fortes da vida mental e psicológica em geral. Admitindo isso, você não tem um e o mesmo corpo que tem na terra, mas por que isso deveria ser importante? Se você sobrevive no caso do transplante cerebral com um corpo numericamente diferente, então por que não nessa situação? Simplesmente ocorre que a transferência da pessoa nesse caso é executada de maneira mais sofisticada, menos confusa. Na verdade, no caso do transplante de cérebro você ainda terminava com um e o mes-

mo cérebro, enquanto nessa situação você adquire um cérebro totalmente novo, tendo o antigo sido vaporizado. Mas de novo, por que isso deveria importar, na medida em que sua psicologia e memórias estão intactas? Além disso, vamos supor que após ter estado algum tempo no planeta distante você se aborreça e deseja viajar. Consequentemente, você entra no equipamento de transferência de estado cerebral de novo muitas vezes, cada vez sendo transportado para um planeta diferente e adquirindo um novo corpo e um novo cérebro a cada parada. Isso se ajusta precisamente ao que Locke disse sobre a identidade pessoal: "Pois, isto sendo a mesma consciência que fez um homem ser ele mesmo para ele mesmo, a identidade pessoal depende apenas disso, que ele seja anexado apenas a uma substância individual, ou possa ser continuado em uma sucessão de várias substâncias".<sup>11</sup>

Parfit também pode lidar muito facilmente com um segundo cenário, em que você entra no cubículo como antes, mas desta vez a máquina funciona mal e não vaporiza seu corpo antes de transmitir a informação ao planeta distante. Consequentemente, acabamos em uma situação tal que a pessoa que entrou no transportador é psicológica e fisicamente contínua com a pessoa que desceu dele na Terra, enquanto a pessoa no planeta distante é psicologicamente, mas não fisicamente, contínua com a pessoa original na Terra. Essa situação é representada diagramaticamente na Figura 11.2.

Qual pessoa tem a melhor alegação para ser você: B ou C (veja Figura 11.2)? Pela mesma lógica que aplicamos no exemplo do cérebro dividido, a afirmação é de que você sobrevive como B e C, ainda que a transitividade da identidade o proíba de ser idêntico com qualquer um. O fato de que um do par, B, seja fisicamente contínuo com você, enquanto C não o é,

irrelevante para o fato de você sobreviver ou não segundo a abordagem de Locke/Parfit. Em matéria de sobrevivência, a continuidade psicológica, e não a física, reina suprema. A teoria de Locke/Parfit também subsidia facilmente a identidade pessoal. Continuando com o esboço que tem sido apresentado da teoria, temos a identidade pessoal apenas naqueles casos em que nenhuma ramificação ocorre. Após a ocorrência da divisão, se um dos hemisférios morre e o outro sobrevive, o sobrevivente será você. Não há outra competição pelo título, que você ganha por desistência. Similarmente, quando o equipamento de transferência de estado cerebral não funciona mal e apenas cria uma única réplica, esta réplica é você, não havendo outro candidato disponível para essa honra. Espera-se que isso tenha esclarecido por que a condição (4) foi incluída na declaração formal de Parfit sobre a identidade pessoal como continuidade psicológica na seção "Lidando com a objeção de Reid".

Como vimos, a abordagem de Locke/Parfit é análoga a tratar você como uma função ou programa que roda no *hardware* do cérebro, sendo o material incorporado estritamente irrelevante para o corpo e o cérebro, assim como a informação em um disquete pode ser transferida para outro disco se o disco original se tornar defeituoso.

Na fantasia descrita, tudo isso é controlado pelo equipamento de transferência de estado cerebral, mas a essa possibilidade pode ser dadã uma guinada teológica. Deus, sendo todo poderoso, pode fazer tudo o que é logicamente possível. Então, o que quer que o transportador possa fazer, Deus pode fazê-lo igualmente bem, se não melhor. É assim que a vida depois da morte é possível, de acordo com essa visão. Com a morte seu corpo e cérebro são destruídos, mas Deus preserva o programa, o fluxo de experiências que exibe encadeamento e continuidade fortes. Ele preserva essas experiências não no *hardware* físico, mas em sua mente. Afinal, não há objeção, como os funcionalistas concordam, à ideia de que um programa poderia não ser igualmente bem instanciado em arranjos não-materiais, bem como em materiais, pois o programa possui uma abstração que significa que ele não pode ser estritamente identificado com os arranjos que o realizam (ver Capítulo 5, seção "Psicofuncionalismo e o conceito de uma pessoa"). Na verdade, Deus – sendo todo bondade assim como todo poderoso – faz melhor do que os equipamentos convencionais de armazenamento de programas. O que ele faz é deletar algumas partes do programa que estão imperfeitas ou danificadas, e, desta maneira, aperfeiçoa-nos para o Céu. Ele é, então, essencialmente um físico

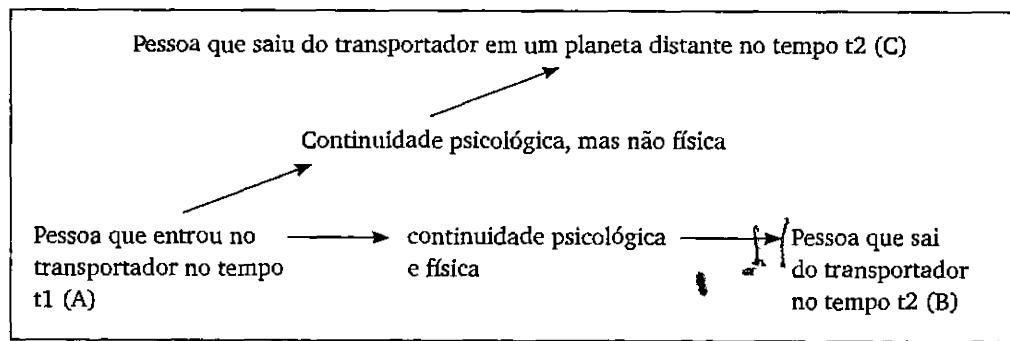


Figura 11.2 Continuidade física e psicológica.

moral, limpando e purificando a alma das pessoas e livrando-as do pecado.

A objeção que se segue pode, porém, ter-lhe ocorrido. Estamos realmente autorizados a concluir que a réplica em um planeta distante é você? Sem alguma espécie de continuidade física, você realmente sobrevive, sendo escaneado na máquina? A máquina não é mais bem descrita como uma câmara de execução, que rapidamente e de modo indolor o desintegra? Nessa visão, a sua réplica no planeta distante é apenas isso, uma mera réplica, uma brilhante falsificação. A sobrevivência dessa réplica não assegura sua sobrevivência mais do que o fato de que seu apêndice será preservado eternamente em um frasco, como Anthony Flew memoravelmente observou uma vez.<sup>12</sup>

Suspeito que as intuições de muitas pessoas concordarão com a conclusão de que o equipamento de transferência de estado cerebral, quando não funciona mal, não é preservador da pessoa. Quando a máquina funciona mal, a pessoa que nela entra sobrevive por um feliz acidente porque seu corpo e, o mais importante, seu cérebro, são deixados inalterados. Mas a completa fundamentação dessa resposta precisa ser articulada, e isto nos leva de volta à continuidade física como um critério de identidade pessoal.

### **A continuidade física revisitada**

Peter Unger argumentou vigorosamente que a continuidade física é necessária para a preservação da identidade pessoal.<sup>13</sup> Ainda que seja impossível aqui fazer justiça ao grande alcance e sutileza da discussão de Unger, sua tese central pode ser colocada de forma direta e simples. Ele admite que a preservação e a continuidade da vida psicológica de uma

pessoa são cruciais para sua sobrevivência e identidade ao longo do tempo. Se seu cérebro fosse misturado de tal forma que todas as suas memórias e outros aspectos psicológicos fossem perdidos, você também seria perdido. Além disso, para que você exista em um tempo futuro, então deve haver, desde agora até o tempo futuro, a contínua existência de suas capacidades mentais básicas. Unger põe grande ênfase na preservação contínua das capacidades psicológicas. Essas capacidades podem continuar a existir, mesmo quando não estão sendo exercidas, ou mesmo durante aqueles períodos em que não podem ser exercidas, como, por exemplo, se seu cérebro estivesse congelado.

Por comparação com Descartes, que pensou que a mente humana precisa ser consciente mesmo durante períodos de sono sem sonhos e no útero da mãe, Unger apenas requer a preservação das capacidades, não seu exercício real durante períodos de consciência. O cérebro, ao contrário do restante do corpo, é vitalmente importante aqui, porque, em comum com outros materialistas, Unger acredita que o cérebro mantém a vida mental. De fato, ele está preparado para ir além, sustentando que cada experiência consciente é um processo físico no cérebro. Mas isso não é uma posição que ele defenda, nem, como ele mesmo aponta, que seja necessária, a fim de elaborar a abordagem da identidade pessoal que ele faz. Observa que poderia bastante satisfatoriamente ser um monista não-reducionista, tanto quanto sua abordagem da identidade pessoal permite.

Unger apresenta sua teoria de que a identidade pessoal é, afinal de contas, baseada na continuidade física do cérebro, não como uma questão de necessidade lógica ou conceitual, mas apenas porque esta é a teoria que os fatos, como nós os conhecemos, irão apoiar.<sup>14</sup> Ele está disposto a levar em conta a possibilidade de

que existam mundos onde a identidade pessoal não seja materialmente baseada, mas esses mundos possíveis não são nosso mundo. No nosso mundo, a identidade pessoal, na prática, depende da preservação continuada da psicologia de uma pessoa por seu cérebro.

Vejamos agora o que Parfit, ao contrário, disse sobre a importância da continuidade física para a identidade pessoal. Parfit diferencia o que ele chama de versões forte, intermediária e fraca do critério psicológico.

De acordo com a versão forte, a causa física da continuidade psicológica deve ser a causa normal, isto é, o cérebro. A versão fraca, no extremo oposto, sustenta que, mesmo se a causa é anormal, a identidade pessoal é preservada. O equipamento de transferência de estado cerebral é supostamente apenas uma tal causa anormal de continuidade psicológica. A ideia aqui é de que, apesar de normalmente o cérebro preservar a continuidade da vida mental, a operação do equipamento de transferência de estado cerebral, ainda que de uma maneira não usual, executa a mesma função, e vai fazê-lo igualmente bem. Como a preservação da continuidade psicológica é executada realmente por um conjunto de arranjos físicos realmente não importa, desde que de uma maneira ou outra isso aconteça. A versão intermediária, como o nome implica, quer o melhor de ambos os mundos. A causa da continuidade psicológica precisa ser tanto o cérebro quanto, ao menos, alguma coisa não muito diferente do cérebro, por exemplo, um cérebro orgânico comum que tenha sido gradualmente transmutado, pela substituição de suas partes por substitutos biônicos para uma estrutura inorgânica, servindo supostamente à função mental executada pelo cérebro original.

Todas as três versões estão de acordo quanto à importância da continuidade

de psicológica para a sobrevivência e a identidade pessoal. Elas diferem apenas com respeito a como isso pode ser obtido satisfatoriamente. Unger discorda vigorosamente da versão fraca. Ele pensa que a descrição do equipamento de escaneamento como um equipamento de transferência de estado cerebral faz uma petição de princípios. O problema com esse tipo de descrição e exemplo de ficção científica – compare os transportadores na *Jornada das Estrelas* – é que ele estimula a crença de que a pessoa é transportada e faz com que sua aceitação pareça inevitável. Nós aceitamos a ideia de que a pessoa é transportada por causa de uma boa história e porque satisfaz nossa imaginação.

Mas seria melhor, propõe Unger, se a máquina escaneadora fosse chamada, de maneira mais neutra, de equipamento de gravação que apenas registra a informação a respeito da posição de suas moléculas. A transmissão dessa informação a um planeta distante, onde ela é usada para construir uma réplica de você, a partir de materiais locais apropriadamente modificados, não apresenta a menor tendência, Unger sustenta, a mostrar que você realmente conseguiu realizar a jornada. Quando você entra no equipamento de gravação, seu corpo e cérebro são vaporizados, e isso é o seu fim. A dissolução de seu cérebro significa o fim da preservação de suas capacidades psicológicas de uma forma que algumas das outras coisas que poderiam ser feitas ao seu cérebro não fazem, por exemplo, a substituição gradual de todas as células orgânicas operantes por substitutos biônicos, como o critério intermediário prevê. Congelar seu cérebro, esquartejá-lo, colocá-lo junto de novo e descongelá-lo, tudo pode ser consistente com a preservação de suas capacidades psicológicas. Mas despedaçar seu cérebro em milhares de células nervosas, mesmo se subsequentemente estas células são

reunidas uma à outra na estrutura original de antes de a catástrofe ter ocorrido, não é igualmente consistente. O cérebro recentemente montado pode sustentar bem a vida mental de uma pessoa, mas esta pessoa não seria você em qualquer sentido, tendo você desaparecido permanentemente no momento em que seu cérebro foi desmembrado em seus componentes.

Se essa última afirmação for correta, então nem mesmo Deus é capaz de ressuscitá-lo por meio do reagrupamento de todas as suas moléculas. Ele pode muito bem recompor todas as partículas de que você era feito, como elas eram quando você era vivo, e ele poderia, portanto, criar uma nova pessoa viva. Mas essa não seria você. A nova pessoa apenas teria sido feita a partir dos materiais com os quais você antes havia sido constituído. A estrutura organizacional que existiu continuamente enquanto você era vivo e que durante aquele tempo fundamentou suas capacidades psicológicas foi perdida. Compare o seguinte caso: uma estátua feita de um bloco de ouro não pode existir sem o ouro, mas não é o mesmo que o bloco. O bloco sobrevive quando fundido, mas a estátua não. Consequentemente, algo é verdadeiro a respeito do bloco que não o é a respeito da estátua e, segundo a lei de Leibniz (veja Capítulo 2, seção "O argumento da dúvida"), isto impede a sua identidade, mesmo que, com uma aparência de paradoxo, não possa haver estátua sem o material que compreende o bloco. Se o material fundido fosse então remodelado na forma da estátua original, e mesmo se por alguma estranha eventualidade todas as moléculas retornassem a suas relações originais umas com as outras, parece que isso não seria uma e a mesma estátua original. Isso não é similar ao caso em que um navio ou uma bicicleta são desmontados e depois remontados.

Esses são objetos compostos, cuja própria natureza consiste em serem montados e permitem que suas partes sejam gradualmente substituídas e que sejam separadas e recompostas, enquanto ainda retêm suas identidades. Não podemos nos permitir a mesma liberdade com respeito a blocos de uma coisa.

Enfatizando a importância da continuidade do cérebro e de sua estrutura na preservação das capacidades psicológicas de uma pessoa, e consequentemente de sua identidade, pode-se pensar se isso nos leva à estranha tese de que uma pessoa não é idêntica a nada, exceto, em última análise, a seu cérebro. Marx chamou esse tipo de posição de materialismo cartesiano. Os teóricos da identidade mente/cérebro parecem estar comprometidos com essa posição, e Thomas Nagel também brincou com a ideia. A estranheza da posição pode ser exibida quando refletimos que, de acordo com o cartesianismo, não podemos dizer que Keith Maslin tem 187cm de altura, ou pesa tantos quilos. Estritamente, temos de dizer que o corpo de Keith Maslin, mas não Keith Maslin, tem essas características físicas, sendo o verdadeiro Keith Maslin uma alma imaterial à qual faltam todos os aspectos físicos.

Sustentar que a pessoa é o cérebro leva a um resultado similarmente estranho. De acordo com essa visão, como para os cartesianos tradicionais, você nunca me encontrou, porque nunca viu meu cérebro. Igualmente, se sou meu cérebro, nós certamente seríamos proibidos de dizer que estou chorando, me apaixonando, ficando zangado, esquecendo de alguém, vendo televisão, lendo um livro ou tocando piano, pois cérebros não podem fazer nenhuma dessas coisas.

Identificar-me com meu cérebro é certamente identificar-me com uma parte de mim, apesar de ser uma parte vitalmen-

te importante, sem a qual eu indubitavelmente não poderia fazer nenhuma das coisas há pouco mencionadas. Felizmente, porém, mesmo se a existência continuada de meu cérebro determinasse minha continuidade e identidade ao longo do tempo, não se segue que preciso ser idêntico a meu cérebro. Retornando ao exemplo da estátua e do bloco de ouro, a identidade das coisas F pode ser fixada pela identidade das coisas G, sem que os Fs sejam idênticos aos Gs, como Brian Garrett aponta. A estátua não pode ser identificada com o bloco pela razão dada anteriormente, mas "as condições de identidade para o bloco fixam as condições de identidade da estátua da seguinte maneira: necessariamente, se o bloco em forma de estátua continua a existir com sua forma bastante inalterada, então a estátua continua a existir".<sup>15</sup> E, podemos acrescentar, se não, não.

Finalizando esta seção, concluo que a continuidade física do cérebro, ou algo muito proximamente equivalente ao cérebro, o que torna possível e sustenta as capacidades psicológicas de uma pessoa continuamente, mesmo se aquelas capacidades não são e não podem ser temporariamente exercidas, é essencial para a identidade pessoal. Shaun agora (lembre-se dele do Capítulo 10) é uma e a mesma pessoa que Shaun então se, e somente se:

1. Shaun agora possui continuidade psicológica no sentido da preservação de suas capacidades psicológicas.
2. Também há continuidade física que, afinal de contas, responsável pela continuidade psicológica. A causa de (1) é uma causa normal, isto é, o cérebro de Shaun, ou talvez algo muito semelhante ao cérebro de Shaun, tal como um

substituto biônico desenvolvido gradualmente. O assim chamado equipamento de transferência de estado cerebral não conta como uma causa normal e aceitável de continuidade psicológica.

3. O desmembramento, que iria fornecer outro candidato igualmente bom ao status de ser Shaun, não ocorreu.

### **Identidade pessoal e não-reducionismo**

Atualmente existem filósofos, ainda que provavelmente apenas uma pequena minoria, que rejeitam as abordagens reducionistas da identidade pessoal em termos seja de continuidade física ou psicológica, ou uma combinação além das duas, esboçadas anteriormente. Para tais filósofos, a identidade pessoal é uma noção conceitualmente primitiva, não-analisável dentro de outros conceitos, tais como da continuidade física do cérebro ou das relações causais entre várias condições ou estados psicológicos, particularmente a memória. A permanência do sujeito da consciência, o ego ou o próprio eu\*, aquele a quem o "eu" supostamente se refere, é representada como um fato subsequente que está acima e além da manutenção de relações mentais ou físicas entre vários estados. Consequentemente, a permanência do próprio eu ou do sujeito de experiências não pode ser reduzida à manutenção de tais relações. No passado essa visão foi destacadamente associada ao filósofo do século XVIII Bispo Joseph Butler que, substituindo uma visão reducionista do próprio eu, realizou um movimento semelhante àquele dos que rejeitaram uma teoria causal da permanência dos objetos

\* N. de T.: No original: *self*.

físicos, a teoria de que a permanência de tais objetos possa ser observada como uma série temporariamente estendida de fatias ou estágios do tempo que são causalmente conectados uns aos outros. A permanência de um objeto físico é algo que está acima e além, e, portanto, não é redutível à obtenção de relações causais entre seus estados físicos em tempos diferentes, ou sua continuidade no espaço.

Ao contrário do que a teoria da continuidade psicológica sugere, uma abordagem não-reducionista da identidade pessoal implica que uma pessoa poderia sofrer de amnésia completa e radical, do tipo mencionado anteriormente, e ainda assim ser uma e a mesma pessoa. Além disso, todas as experiências passadas e antigos traços psicológicos poderiam ser varridos, e um novo conjunto de experiências e características instalado, e ainda assim ela seria uma e a mesma consciência, um e o mesmo eu, ao longo de todas as mudanças. Se for verdade que a identidade pessoal consiste em um fato subsequente acima e além do encadeamento mental, então nenhuma contradição pode possivelmente resultar da suposição de que uma pessoa sobreviverá às drásticas mudanças à sua vida psicológica que estamos considerando. Será uma e a mesma pessoa após as mudanças como fora antes que as drásticas alterações tivessem sido implementadas.

Em apoio a tal afirmação, Richard Swinburne escreveu:

Muitos religiosos têm levado a sério lendas sobre pessoas passando através das águas do Lethe (um rio cujas águas fazem uma pessoa esquecer tudo sobre sua vida prévia), adquirindo depois um novo corpo. Outros que escutaram essas lendas podem não ter acreditado que fossem verdadeiras; mas usualmente as compreendem, e (a menos que influenciados por

um dogma filosófico) não têm suspeitado que elas envolvam contradição.<sup>16</sup>

Igualmente, de acordo com a teoria que observa a identidade pessoal como primitiva e não-analisável, a identidade de uma pessoa nada tem a ver com a identidade de seu corpo, ou de qualquer parte dele. Afinal, o que ocorre ou não ocorre ao corpo, incluindo o cérebro, não tem efeito sobre a identidade da pessoa, ou relevância para ela.

Nesse meio tempo, deve estar claro que, de fato, a abordagem não-reducionista da identidade pessoal resulta na adoção do dualismo de substâncias. O próprio eu, o sujeito de consciência, não pode ser identificado com um objeto material, tal como o corpo ou o cérebro, nem como consistindo em uma sucessão de estados físicos causalmente relacionados um ao outro. Mas nem o próprio eu é redutível à série de estados mentais que possui durante o período de sua existência. Ele é aquilo que possui os estados mentais em questão, mas não pode ser identificado com os estados que possui. Como o próprio eu é esta coisa singular, este sujeito unitário de experiências que persiste e continua o mesmo ao longo de mudanças, até mesmo muito radicais, em sua vida psicológica, e uma vez que, em acréscimo, sua identidade não tem absolutamente nada a ver com a persistência e a identidade das coisas materiais, tais como o corpo e o cérebro, parece se seguir que o próprio eu tem de ser concebido como, *p ace* cartesianismo, uma substância simples, logicamente indivisível.

Essa teoria, se correta, tem importantes consequências para o experimento mental conduzido, anteriormente a respeito da divisão do cérebro em dois hemisférios, cada um dos quais sendo então transplantado. Isso foi entendido como demonstrando a possibilidade da sobre-

vivência sem identidade, constituindo a identidade pessoal o caso especial em que o não-desmembramento ocorre, como é efetivamente o caso quando um dos hemisférios transplantados não sobrevive à operação. O hemisfério sobrevivente é, para todos os efeitos e propósitos, identificável com a pessoa original, não havendo algum outro candidato adequado disponível.

Para o dualismo, porém, nenhuma destas irá ocorrer. Diferentemente dos casos em que bulbos de tulipas, amebas ou rios se dividem, e então em algum sentido sobrevivem como as partes de novas entidades, substâncias espirituais simples não podem se dividir. Tais almas não são feitas de partes menores que se pareçam com almas; portanto, embora quando o cérebro é dividido parte dele vá para a direita e parte para a esquerda, não faz sentido dizer que metade da alma vá para a esquerda e metade para a direita. Então, no caso das almas, somos proibidos de ter sobrevivência sem identidade. Se a alma sobrevive à divisão do cérebro (e há toda razão para supor que iria, pois ela não depende para sua existência de estados de coisas físicos), então há apenas três possibilidades quanto ao que ocorre a ela em relação aos dois organismos físicos resultantes, cujos crânios contêm os hemisférios esquerdo e direito do cérebro original. Primeiramente, a alma vai para onde o hemisfério direito vai. Em segundo lugar, ela vai para onde o hemisfério esquerdo vai. Em terceiro lugar, ela não vai para hemisfério algum. Se essa terceira possibilidade fosse realizada, então o resultado seria a criação de dois organismos sem alma – de fato, zumbis –, uma vez que nenhuma mentalidade seria associada a qualquer das entidades. Haveria um corpo vivo funcionando contendo um cérebro funcionando; contudo, porque a mentalidade não é superveniente sobre

processos cerebrais, nem idêntica a eles, esses corpos careceriam de qualquer vestígio de uma vida mental. Não haveria, por assim dizer, ninguém em casa em ambos os casos.

Contudo, se descontamos essa terceira possibilidade, então a pessoa, como uma alma, sobrevive, seja em conexão com o hemisfério direito ou com o esquerdo, mas não com ambos simultaneamente. De certo modo, esse é um resultado satisfatório, porque apóia uma intuição sobre a identidade pessoal que é difícil de abalar, a saber, que isto é tudo ou nada e não admite graus, ao contrário do que as teorias da continuidade psicológica e física parecem implicar. Poderia haver outra razão pela qual esse resultado pode ser pensado como sendo satisfatório. Imagine que você mesmo está prestes a ter seu cérebro dividido, porque seu corpo original está tomado por um câncer inoperável. Você está preocupado a respeito de sua sobrevivência, mas é tranquilizado pelo cirurgião, que leu Parfit, e que lhe diz que o resultado é melhor que a morte, porque você sobreviverá parcialmente como duas pessoas. A visão singular que você tem agora do mundo será dividida em duas, e então você será capaz de ver o mundo de duas perspectivas diferentes, e ambas, em algum sentido, serão suas. Após a divisão, você não mais será capaz de saber, sem observação e inferência, o que está acontecendo na sua outra mente, mas, de alguma maneira, ainda vai estar por aí como dois centros de consciência e experiência distintos e independentes.

Porém, você pode bem hesitar diante da descrição do que vai ocorrer a você, reclamando que isso cheira a incoerência. Como o seu presente ponto de vista singular do mundo poderia se tornar dois pontos de vista no futuro? Como pode o que parece ser essencialmente um ponto de-vista singular vir a tornar-se dois pon-

tos de vista? Uma variação desse tema é fornecida pela imagem mental de Bernard Williams sobre um cirurgião louco que diz que, depois de seus hemisférios serem transplantados, uma das pessoas resultantes será horrivelmente torturada, enquanto a outra será libertada e premiada com um milhão de libras. Como você deve julgar tal perspectiva? O que você quer saber é se será azarado e receberá a tortura, ou será sortudo e ganhará o milhão de libras. Mas, para aqueles que acreditam na sobrevivência parcial sem identidade, essa difícil escolha de resultados alternativos não se aplica. Em vez disso, somos forçados a dizer que como uma das pessoas você será torturado e como a outra pessoa você ganhará um milhão de libras, mas não há uma pessoa singular que tanto receba a tortura quanto o milhão de libras. Então nenhuma resposta claramente determinada emerge quanto ao que irá ocorrer a você, o único sujeito de experiências, tal como você se encontra neste momento, no futuro, após os transplantes terem sido realizados.

Ao contrário, se você é uma alma, deve haver uma resposta determinada à questão do que lhe ocorrerá após a operação. Ou sua alma irá onde o hemisfério direito for, ou irá onde o hemisfério esquerdo for, e um dos indivíduos resultantes receberá a tortura, e o outro, o milhão de libras. Então, segundo a tese dualista, podemos estar certos de que um único indivíduo vai receber seja a tortura, seja o milhão de libras. Há, em outras palavras, uma resposta definida quanto a quem por fim receberá a tortura e quem receberá o prêmio. Antes de a operação ocorrer, a pessoa em questão pode apenas esperar que, como uma alma, ela seguirá o hemisfério que ajuda a constituir a pessoa que é premiada e não torturada. Swinburne afirma que nem leis naturais nem raciocínios filosóficos podem determinar a qual

hemisfério transplantado o espírito irá por fim se unir. A designação de um hemisfério particular para um espírito particular é, sustenta Swinburne, ou deixada a Deus ou à sorte.

Porém, de acordo com a teoria da mente monista e não-reducionista adotada por este livro, concordo com a observação de J. L. Mackie de que "uma visão materialista do pensador" é menos controversa que uma visão materialista dos pensamentos.<sup>17</sup> Portanto, enquanto há razões muito fortes para não tentar reduzir estados mentais a estados cerebrais, razões igualmente fortes puxam na direção oposta quando se trata da natureza do sujeito de consciência. É mais plausível ver as identidades das pessoas como dependendo dessas estruturas físicas do que sustentar e tornar possível a continuidade de várias espécies de capacidade psicológica. Além disso, como vimos no Capítulo 2, o problema de como espíritos incorpóreos devem ser individuados e identificados significa que o dualismo de substâncias é, por fim, uma abordagem profundamente não atraente sobre a natureza das pessoas, e por isso uma visão reducionista da identidade pessoal é muito mais provável de ser verdadeira.

### Questões para reflexão

1. Qual o grau de plausibilidade da abordagem de Parfit sobre a identidade pessoal em termos de continuidade psicológica?
2. Em *Jornada nas Estrelas*, quando o Capitão Picard é transportado de ida e volta à superfície de um planeta, é realmente Picard quem sai do transportador ou uma brilhante réplica?
3. Que problemas filosóficos surgem quando nos esforçamos em estabe-

- lecer nossa identidade contínua ao longo do tempo?
- 4. É possível que você sobreviva como cada uma de duas pessoas diferentes?
  - 5. Se você for sofrer um apagamento completo da memória, deve preocupar-se com o que ocorrerá ao seu corpo posteriormente?

### Sugestões para leituras aprofundadas

- PARFIT, D. *Reasons and Persons*. Oxford, Clarendon Press, 1984. Este é um texto chave no debate moderno, sendo vivaz, original, divertido e altamente recomendado.
- DANCY, J., ed. *Reading Parfit*. Oxford, Blackwell, 1997 contém uma importante coleção de artigos sobre Parfit.
- UNGER, P. *Identity, Consciousness, and Value*. Oxford, Oxford University Press, 1990 é um livro importan-

te que elabora um caso altamente persuasivo para a abordagem da continuidade física da identidade pessoal.

- GARRETT, B. *Personal Identity and Consciousness*. London and New York, Routledge, 1998 reúne as abordagens física e psicológica da identidade pessoal, e é útil pela visão geral que fornece, bem como pela clareza de suas discussões.
- WILKES, K. V. *Real People: Personal Identity without Thought-Experiments*. Oxford, Clarendon Press, 1988 é um livro que contém discussões intrigantes de múltiplas personalidades e dos efeitos da comissurotomia. Também é útil por sua discussão da consciência e da imagem aristotélica da alma.
- WILKES, K. V. "Psuche versus mind". in: NUSSBAUM, M. e RORTY, A. O., eds. *Essays on Aristotle's De Anima*. Oxford, Clarendon Press, 1992. Esse artigo também deve ser visto.

### NOTAS

1. LOCKE, J. *An Essay concerning Human Understanding*. London, Everyman's Library, 1961. v. II, c. XXVIII, seção 10, ênfase minha.

2. Ibid.

3. PARFIT, D. *Reasons and Persons*. Oxford, Clarendon Press, 1994.

4. Ibid., p. 207.

5. BUTLER, J. "Of Personal Identity". in: BERNARD, J. H., ed. *The Works of Bishop Butler*. London, s. ed., 1990. v. II

6. McGINN, C. *The Character of Mind*. Oxford, Oxford University Press, 1992, p. 112:

Visto que é perfeitamente possível a existência de relações de dependência causal entre estados mentais de pessoas distintas – como quan-

do uma pessoa vem a ter uma crença como resultado do que outra disse –, nós claramente precisamos de algumas restrições a respeito de quais relações causais têm condições de vincular estados da mesma pessoa. E é de fato surpreendentemente difícil suprir quaisquer restrições que não importem circularidade e ainda tenham uma chance de funcionar.

7. PARFIT, *Reasons and Persons*, p. 251.

8. O'HEAR, A. *What Philosophy Is*. Atlantic Highlands, NJ, Humanities Press International, 1985, p. 252.

9. PARFIT, *Reasons and Persons*, p. 281-2.

10. Ibid., p. 282

11. LOCKE, J. *Treatise concerning Human Understanding* (várias edições), v. II, c. XXVIII, seção 10, ênfase minha.

12. FLEW, A. G. N. *Body, Mind and Death*. New York and London, Macmillan, 1964, p. 5.
13. UNGER, P. *Identity, Consciousness, and Value*. Oxford, Oxford University Press, 1990.
14. Ibid., p. 102. Compare John Mackie:

Eu diria que é uma questão empírica o que torna a coconsciência possível [e por implicação, a identidade pessoal], assim como seria uma questão empírica saber qual constituição interna gera as propriedades prontamente observadas do ouro. E se essa é uma questão empírica, agora sabemos ao menos o esboço da resposta: o que torna a coconsciência possível é a estrutura do sistema nervoso central e a permanência dessa estrutura ao longo do tempo... O conceito de identidade pessoal, tal como o estamos agora interpretando, não é o conceito de uma continuidade corporal, mas é o conceito de algo que vem a ser a continuidade de certa parte do corpo [o cérebro]. É concebível que a identidade pessoal deva ter sido a permanência de alguma substância imaterial, pois poderia ter sido sobre isso que a possibilidade

de uma co-consciência de experiências dependeu. Não temos razão alguma para supor que isso aconteça: mas é uma verdade empírica, e não uma verdade analítica, a de que não sobrevivemos à morte corporal.

Essa interpretação do conceito de identidade pessoal é certamente muito análoga à visão de Armstrong da mente como por definição a causa interna do comportamento, e, por isto, contingentemente, o sistema nervoso central. Mas essa é uma resposta mais satisfatória a essa questão, porque o complicado tema da identidade de propriedades aqui não surge. Uma visão materialista do pensador é menos controversa que uma visão materialista dos pensamentos. (MACKIE, J. *Problems from Locke*. Oxford, Clarendon Press, 1976, p. 200-2.)

15. GARRATT, B. *Personal identity and Self-Consciousness*. London, Routledge, 1999, p. 10.
16. SWINBURNE, R. *Personal Identity*. Blackwell, Oxford, 1994, p. 25.
17. MACKIE, J. *Problems from Locke*, p. 202.

# GLOSSÁRIO

## *A fortiori*

Latim: "pelo mais forte". É usado para querer dizer "tudo o mais" ou "ainda mais". Por exemplo, no coração do argumento cosmológico para a existência de Deus está o argumento ontológico, e como este último argumento é profundamente imperfeito, então, *a fortiori*, o argumento cosmológico está profundamente roto.

## *Ampliativo*

Um termo que é aplicado a **argumentos** cujas conclusões contêm informações adicionais, não encontradas em suas **premissas**. Consequentemente, amplia o que está para ser encontrado nessas premissas. **Argumentos indutivos** fornecem o caso modelo. Por exemplo, a conclusão de que o sol vai continuar a surgir porque ele sempre surgiu no passado vai além de nossa evidência presente para esta crença.

## *A posteriori*

O que se conhece tendo por base a experiência, ou depois dela. Por exemplo, causas apenas podem ser conhecidas por meio da experiência, e não independentemente dela. Proposições *a posteriori* são aquelas cuja verdade ou falsidade pode apenas ser conhecida tendo por base a experiência.

## *A priori*

O que se conhece como sendo verdadeiro independentemente da experiência, e, portanto, anterior a ela. Por exemplo, as proposições da matemática e da geometria, tais como "2 + 2 = 4" e "Todos os triângulos euclídeanos têm ângulos somando 180 graus", são conhecidas

como sendo verdadeiras em virtude do significado das palavras nelas contidas. Não podem ser confirmadas ou infirmadas por observação e experimentação.

## *Argumento*

Na filosofia, um argumento não é uma disputa, mas uma tentativa de se mover racional ou logicamente a partir de sentenças chamadas de "**premissas**", a fim de estabelecer a verdade de outra sentença, chamada de "conclusão". Há duas espécies de argumento, dedutivo e indutivo. Em um **argumento dedutivo** que seja válido, a conclusão se segue logicamente das premissas, de tal maneira que, se as premissas forem asseridas e a conclusão negada, resulta uma contradição. As conclusões de **argumentos indutivos** não se seguem logicamente das premissas, mas vão além delas no sentido explicado no verbete "**ampliativo**".

## *Atomístico*

Estados mentais intencionais, tais como crenças e pensamentos, são **holísticos**. Isso significa dizer que eles pressupõem e acarretam toda uma rede de outros estados intencionais e não podem ocorrer isoladamente. A crença de que o lixeiro não recolheu o lixo, por exemplo, acarreta todas as espécies de crenças, tais como a relativa a que dia da semana é hoje, à existência de um serviço de coleta, e assim por diante. Por outro lado, estados não-intencionais, como sensações, são atomísticos no sentido de que eles podem ocorrer por si mesmos e não conduzir a qualquer implicação sobre a existência de qualquer outro estado da mente.

### **Behaviorismo analítico**

Uma forma de reducionismo que afirma que todo o falar sobre estados mentais pode ser expresso, mais perspicazmente e sem perda de sentido, ao falar sobre padrões reais e possíveis de comportamento.

### **Bicondicional**

Uma sentença bicondicional é aquela que afirma **condições necessárias e suficientes** logicamente e adota a forma “se *p* então *q*, e se *q*, então *p*”, onde *p* e *q* estão para proposições. Isso é frequentemente abreviado a “*p* se, e somente se, *q*” ou “*p* sse *q*”. Por exemplo, “se Martin é um filho único, então ele não tem irmãos e irmãs, e se Martin não tem irmãos e irmãs, então ele é filho único”. Isso pode não soar muito esclarecedor, mas na filosofia da mente uma tentativa tem sido feita para encontrar verdadeiros bicondicionais da forma “*x* é um estado mental *m* se, e somente se, seu cérebro está no estado físico *f*”. Se essa última sentença fosse verdadeira, então estados mentais seriam uma e mesma classe de fenômenos que estados cerebrais, e o mental efetivamente teria sido reduzido ao físico.

### **Cartesianismo**

Em um sentido amplo, é o conjunto de doutrinas e crenças associadas a René Descartes (1596-1650), mas, no contexto da filosofia da mente, é a doutrina **dualista** de que os seres humanos são compostos de duas espécies radicalmente diferentes de coisas, corpo e alma, capazes de existir independentemente uma da outra, mas unidas na vida. A alma é não-física e não possui extensão, enquanto o corpo possui extensão, mas não possui pensamento.

### **Causação**

A causação opera onde um evento suscita, exige ou provoca outro evento. Como o conceito deve, em última análise, ser analisado e compreendido, é uma questão crucial na filosofia. Hume propôs uma célebre análise da causação que a vê como a conjunção constante de eventos – um tipo de evento regularmente seguindo outro – sem conexão discernível ou “cola” para ligar um ao outro. Hume também sugeriu que “A causa B” pode ser entendido como dizendo que se A não ocorresse, então B não aconteceria.

### **Causação psicofísica**

Causação da mente ao corpo. Em outras palavras, estados mentais agem como causas para suscitar o comportamento corporal. Ainda há um enigma no que concerne ao que estados mentais precisam ser para que possam fazer isso. **Materialistas** e **fisicalistas** sustentam abertamente que estados mentais precisam, em última análise, ser físicos em sua natureza, mas há sérias dificuldades em reduzir o mental ao físico, pois isso deixa de fora a consciência e os qualia.

### **Cognição**

A cognição se refere às operações do intelecto e do entendimento por meio das quais chegamos ao conhecimento e às crenças.

### **Computação/computacional**

As operações realizadas por um computador para executar uma seqüência de operações definidas a fim de, por exemplo, encontrar as soluções para problemas aritméticos ou matemáticos. O computador realiza uma série de passos distintos, chamados de algoritmos, que não exigem qualquer *insight* ou inteligência para serem executados. A teoria computacional da mente afirma que a mente é mais bem compreendida como um programa de computador rodando no mecanismo úmido (*wetware*) do cérebro. Veja isso em conexão com o **funcionalismo**.

### **Conação/conativo**

O aspecto da mente que tem a ver com fazer tentativas, agir e querer.

### **Condições necessárias e suficientes**

Isso é mais bem explicado por meio de um exemplo. Uma condição necessária de ser um solteiro é que a pessoa em questão seja homem. Mas isso não é suficiente, porque uma criança recém-nascida do sexo masculino não é um solteiro. O que ainda precisa ser acrescentado a isso é que, em primeiro lugar, a pessoa em questão está com idade adequada para casar-se e, em segundo lugar, que ela não tenha se casado antes. Essas duas condições, associadas à masculinidade, são logicamente necessárias e suficientes para abranger o estado de estar solteiro. Considere agora a noção de condição sufi-

ciente. Uma condição suficiente para ser um cidadão britânico é que alguém tenha nascido de pais britânicos na Grã-Bretanha. Mas isso não é necessário, pois a cidadania também pode ser adquirida de outras maneiras, como pelo casamento. Veja também **bicondicional**.

#### **Condicional contrafatual**

Um condicional contrafatual é uma sentença que toma a forma “Se *x* ocorresse, então *y* ocorreria” ou “se *x* não tivesse ocorrido, então *y* não teria ocorrido”. Todas as afirmações causais, incluindo as leis da natureza, implicam contrafatuais, e não são meras afirmações de regularidades. “Riscar o fósforo na caixa causou sua entrada em combustão” implica que “se o fósforo não tivesse sido riscado na caixa, tudo sendo igual, ele não teria entrado em combustão” e “se o fósforo fosse riscado na caixa, tudo sendo igual, então ele entraria em combustão”. O objetivo da cláusula “tudo sendo igual” é impedir a operação de causas interferentes, por exemplo, os fósforos estarem molhados, o que iria impedir o resultado usual de riscar o fósforo.

#### **DAL (dispositivo para a aquisição de linguagem) – LAD (language acquisition device)**

O linguista americano Noam Chomsky afirmou que a teoria do aprendizado encontrada na psicologia estímulo/resposta não pode explicar como as crianças adquirem a linguagem tão rapidamente, nem sua habilidade de entender e gerar uma variedade indefinida de sentenças e construções que elas nunca encontraram antes. Para explicar isso ele postulou um DAL, um dispositivo para a aquisição de linguagem, encarnado na maquinaria neuronal do cérebro, que incorpora uma gramática universal inata, não aprendida pela experiência, que fornece as regras que as crianças precisam para reconhecer e classificar os diferentes exemplos de expressões que encontram. Dessa maneira, a criança rapidamente aprende quais expressões são significativas, e quais devem ser descartadas.

#### **Dedução/argumento dedutivo**

Em um **argumento** que é dedutivamente válido, a conclusão é logicamente implicada

pelas **premissas**, de tal forma que propor as premissas, mas negar a conclusão, leva a uma contradição. Isso porque a informação contida na conclusão já está presente, implicitamente, nas premissas.

#### **Definição ostensiva**

Uma definição ostensiva é aquela que tem lugar ao apontar. Nem todas as palavras podem ser definidas puramente em termos de outras palavras. É preciso chegar a um ponto em que o termo tem de ser definido por apontar um estado de coisas não-lingüístico ao qual ele pode ser aplicado.

#### **Descrição agentiva**

Uma descrição agentiva é aquela que caracteriza o comportamento de alguém utilizando termos mentais, incluindo aqueles que descrevem o comportamento como uma ação praticada pelo agente. Uma descrição não-agentiva, ao contrário, descreve o comportamento utilizando apenas conceitos retirados da física, e não leva a qualquer implicação de que a mentalidade esteja presente. Por exemplo, “Ele ergueu seu braço” seria uma descrição agentiva porque ela atribui uma ação, praticada intencionalmente e escolhida por uma razão. Ao contrário, a descrição “Um braço humano parece se mover por regiões intervinientes do espaço de uma maneira tal e tal e ao longo de tal e tal período de tempo” não conduz a nenhuma implicação sobre se uma ação foi praticada ou não.

#### **Disjunção/disjuntivo**

Uma sentença disjuntiva é do tipo “ou/ou” que afirma alternativas. Pode haver afirmações disjuntivas inclusivas – por exemplo, você pode tanto beber cerveja quanto comer bolo, ou ambos – e disjunções exclusivas – por exemplo, você pode tanto beber cerveja quanto comer bolo, mas não ambos.

#### **Disposição/disposicional**

Uma propriedade disposicional de uma coisa afirma como essa coisa se comportaria se certas condições fossem satisfeitas. “Solubilidade” é uma propriedade disposicional do açúcar; isso significa dizer que quando imerso na água, por exemplo, ele se dissolverá. Similarmente,

uma pessoa que tem uma disposição à irritação não precisa estar realmente fumegando e reclamando, mas é propenso a assim proceder em circunstâncias que, por outro lado, levariam pessoas de temperamento mais sereno a se manter tranquilas. Veja também **ocorrente**.

### **Dualismo/dualista**

O dualismo toma duas formas: dualismo de substâncias e dualismo de propriedades. De acordo com o dualismo de substâncias, todos os seres humanos são compostos de duas espécies radicalmente diferentes de substâncias ou coisas: mentes não-físicas e corpos físicos, cada uma capaz de existir independentemente da outra. O dualismo de propriedades postula apenas substâncias físicas, mas afirma que elas possuem duas espécies diferentes de propriedades: propriedades físicas e propriedades mentais não-físicas, tais como a consciência. Veja também **monismo não-redutivo**.

### **Eliminativismo**

Associado particularmente à obra de escritores tais como Richard Rorty e Paul e Patricia Churchland, o eliminativismo não só nega que haja coisas como estados mentais, mas advoga a eliminação de todos os termos psicológicos e mentais do nosso vocabulário, recomendando que sejam substituídos por termos extraídos das ciências, especialmente das ciências computacional e cognitiva.

### **Epifenomenalismo**

Epi é o termo grego para “em cima de”, então o epifenomenalismo significa literalmente “um fenômeno ocorrendo em cima de outro”. A ideia básica é de que um fenômeno produz outro que é dependente dele, assim como uma luz adequadamente colocada e um objeto projetam uma sombra em uma parede. O comportamento da sombra é dependente do comportamento da luz e do objeto, e não pode variar independentemente da forma como eles se comportam. A sombra não possui a capacidade de produzir qualquer alteração na luz e no objeto. De maneira semelhante, tem sido sustentado que processos físicos no cérebro produzem e sustentam estados mentais, mas que estes estados não podem ter qualquer efeito nos processos cerebrais dos quais eles de-

pendem, e são causalmente impotentes para suscitar quaisquer mudanças físicas.

### **Epistemologia/epistemológico**

Epistemologia é a teoria do conhecimento, e ocupa-se com temas como o que é conhecimento, como conhecemos, e se podemos conhecer algo.

### **Estados mentais não-intencionais**

São estados mentais como as sensações e os estados de ânimo que não são relativos a um conteúdo, ou a ele dirigidos, de tal modo que não representam outro estado possível de coisas, diferentemente dos estados intencionais. Estados não-intencionais são **não-holísticos**.

### **Falácia do homúnculo**

Homúnculo, do latim *Homunculus* (plural: *homunculi*), literalmente significa “homem pequeno” ou “anão”, e veio a denotar o erro na filosofia da mente de atribuir a algum agente interno, seja o cérebro ou a alma, exatamente os mesmos aspectos do agente que foi invocado para explicar em primeiro lugar. Por exemplo, na seção infantil “The Funday Times”, do *The Sunday Times*, havia um desenho animado chamado “The Numbskulls” (“Os Patetas”). Os Patetas são um grupo de pequenos homens que vivem dentro do crânio de Uncle (Tio) e explicavam como ele era capaz de fazer várias coisas, tais como ver e pensar. Um pequeno homem, por exemplo, olhou para fora através dos olhos de Uncle e o dirigiu pelo mundo. Outro pequeno homem no cérebro sugeriu vários pensamentos a ele. O problema, é claro: como esses pequenos homens, por sua vez, vêem e pensam? Eles possuem homens ainda menores dentro deles? Claramente, existe a ameaça do vício do regresso ao infinito, e nesse caso a explicação não leva a nada. Descartes cometeu a falácia do homúnculo quando explicou o que ver significa para nós por meio da postulação de que a pessoa, concebida como uma alma fantasmagórica, vê imagens do mundo externo projetadas na glândula pineal no interior do cérebro.

### **Fenomenalismo**

Fenomenalismo é uma espécie de **redução analítica** e tenta exprimir todo o falar

sobre objetos físicos em termos de experiências reais e possíveis.

### **Fenômenos físicos**

É extremamente difícil definir os fenômenos físicos, mas alguém pode esperar deles (1) que sejam publicamente disponíveis em teoria a todos os sujeitos de experiência; (2) que sejam localizados no espaço, e nele ocupem uma extensão ou área; (3) que existam independentemente de todos os observadores; (4) que frequentemente sejam mensuráveis e quantificáveis. Objetos naturais, artefatos humanos, fenômenos, tais como a luz, a gravidade, o magnetismo, a massa, a solidez, e assim por diante, são exemplos típicos de fenômenos físicos.

### **Fenômenos mentais**

Fenômenos mentais compreendem estados da mente, incluindo acontecimentos e processos mentais. Um excelente exemplo de estado mental é a consciência, e a consciência, de diferentes maneiras, parece ser constitutiva da existência não apenas de estados não-intencionais, como sensações e estados de ânimo, mas também da vasta panóplia de estados intencionais, como crenças, pensamentos, desejos e emoções.

### **Fiscalismo**

O fiscalismo está muito próximo do materialismo e sustenta que tudo o que existe é físico. Frequentemente faz a afirmação adicional de que as únicas explicações autênticas de eventos são as explicações físicas, e os únicos conceitos adequados são os conceitos físicos.

### **Funcionalismo/psicofuncionalismo**

O funcionalismo concebe a mente como uma função que roda no *hardware* do cérebro, por meio de *inputs* sensoriais que são convertidos em *outputs* comportamentais. O funcionalismo metafísico fornece uma análise filosófica puramente conceitual dos *inputs* e *outputs* que constituem um estado mental – digamos, o de dor. O psicofuncionalismo investiga empiricamente quais são os reais mecanismos neurais que estão envolvidos na incorporação e descarga da função. Funções são abstratas em sua natureza e não podem ser estritamente identificadas com os arranjos que ocorrem para in-

corporá-las, quer estes compreendam estados de coisas físicas, ou mesmo conceitivamente não-físicos.

### **Holismo/holístico**

Holismo é o traço pelo qual estados mentais **intencionais** pressupõem logicamente a existência de outros estados intencionais, consequentemente compreendendo uma rede ou sistema de apoio. Diferentemente das sensações que são não-intencionais, estados intencionais não podem existir em isolamento uns dos outros. Uma analogia possível é com uma peça de quebra-cabeça, que pressupõe a existência de outras peças do quebra-cabeça. Veja também **atomístico**.

### **Identidade pessoal**

A questão central é: o que torna uma pessoa X no momento presente uma e a mesma que uma pessoa Y no passado? O que é necessário é que haja **condições necessárias e suficientes** da identidade numérica da pessoa ao longo do tempo.

### **Indução/argumento indutivo**

Argumento indutivo é aquele em que as **premissas** não implicam logicamente a conclusão, porque a conclusão contém informações que vão além das encontradas nas premissas. Tais **argumentos** são desse modo ditos **ampliativos**. A forma típica de um argumento indutivo é: "Todos os Xs observados até então têm sido Ys, portanto, todos os Xs que existem são Ys". Usamos essa espécie de raciocínio milhares de vezes diariamente, e é claro que se não o fizéssemos haveria pouco, além dos dados fornecidos pela experiência atual, que poderíamos afirmar saber. O problema sobre se a indução pode ser racionalmente justificada, a respeito do qual Hume (1711-1776) foi o primeiro a nos alertar, permanece um tema muito debatido até hoje.

### **Intencionalidade/intencional**

Intencionalidade é o aspecto por meio do qual muitos estados mentais possuem um conteúdo representacional. Alguns estados intencionais, tais como crenças, pretendem representar como o mundo é realmente. Se o mundo é como uma crença o representa como



sendo, a crença é verdadeira; de outra forma ela é falsa. Outros estados intencionais, como desejos e intenções, representam como um agente quer que o mundo seja. Desejos e intenções que não conseguem suscitar os estados de coisas que pretendem são frustrados ou insatisfeitos. Veja também **sobreidade**.

#### **Intencionalidade intrínseca e derivada**

Muitos filósofos, dos quais John Searle seria um excelente exemplo, têm afirmado que estados intencionais são intrínsecos aos cérebros, e não derivados de outros sistemas intencionais. Programas de computador, ao contrário, não são intrinsecamente relativos a outros estados de coisas, e não possuem uma intencionalidade genuinamente não-derivada. O conteúdo de tais programas existe apenas porque estamos aqui para interpretar os programas. Em um mundo desprovido de mentes, programas de computador não possuiriam um conteúdo, bem como linguagem, livros e símbolos em geral. Dois pedaços de madeira, em forma de cruz, não significam nada por si mesmos. Nós atribuímos um sentido a tal estrutura, transformando-a, pois, no símbolo cristão para a redenção da humanidade por meio de Cristo.

#### **Lei de Leibniz**

A lei de Leibniz, formulada pelo grande filósofo alemão Gottfried Leibniz (1646-1716), também é conhecida como a “identidade dos indiscerníveis”. Ela afirma que se um objeto X é uma e a mesma coisa que um objeto Y, então todas as propriedades de X devem ser as mesmas que todas as propriedades de Y. Consequentemente, se X possui ao menos uma propriedade ausente em Y, e vice-versa, então X e Y não podem ser numericamente idênticos, isto é, um e o mesmo. A lei de Leibniz é uma ferramenta útil para avaliar afirmações sobre a identidade. Por exemplo, se estados mentais são idênticos a estados cerebrais, então todas as propriedades de estados mentais precisam ser possuídas por estados cerebrais, e vice-versa. Se apenas uma propriedade pode ser encontrada que seja possuída por um estado mental, mas não por um estado cerebral, e vice-versa, então os estados mentais não podem ser estados cerebrais. Há exceções à lei de

Leibniz. Ela não funcionará, por exemplo, com verbos que denotam estados mentais intencionais. Pode ser verdadeiro que Lois Lane esteja imaginando beijar o Super-Homem e que não esteja imaginando beijar Clark Kent (então algo é verdade a respeito do Super-Homem, mas não de Clark Kent), mas isso não pode estabelecer que o Super-Homem não seja Clark Kent (nós sabemos muito bem que ele é).

#### **Materialismo/fiscalismo**

De acordo com o materialismo, tudo o que realmente existe são objetos físicos e fenômenos físicos. Estados mentais, em última análise, não são nada além de estados físicos de coisas, incluindo as propriedades.

#### **Mereologia**

A mereologia lida com a lógica das relações das partes com o todo, do qual elas são constitutivas.

#### **Metafísica**

A metafísica suscita questões sobre a natureza da existência que está além do escopo das ciências. Questões típicas na metafísica são: Qual é a natureza da causação? O que é a mente, e como ela difere e se relaciona com o corpo? O que constitui a identidade de uma pessoa ao longo do tempo? Qual a diferença entre as coisas e as suas propriedades?

#### **Monismo**

O monismo é a concepção de que apenas um único tipo de coisa existe. O monismo materialista sustenta que as espécies de coisas que existem são físicas em sua natureza. O monismo imaterialista sustenta o oposto: apenas substâncias espirituais, isso é, almas não-físicas e suas propriedades existem.

#### **Monismo não-redutivo**

O monismo não-redutivo sustenta que apenas entidades físicas existem, mas essas possuem tanto propriedades físicas quanto não-físicas, e as últimas resistem à redução das primeiras. O **dualismo** de propriedades é um exemplo de monismo não-redutivo.

#### **Naturalismo**

A concepção de que tudo o que existe é parte do mundo natural e explicável em ter-

mos extraídos das ciências naturais, como a física, a química e a biologia.

### **Naturalismo biológico**

Esse é um termo associado à filosofia da mente do filósofo americano John Searle. Estados mentais constituem uma classe de fenômenos naturais, a saber, fenômenos biológicos. Eles não estão fora do âmbito da natureza, em um mundo sobrenatural ou não-natural, mas fazem tanto parte de nossa história biológica quanto processos como a digestão e a secreção da bile pelo fígado. Há uma implicação de que estados mentais podem ser associados apenas a formas de vida orgânica, e não a artefatos inorgânicos, como computadores ou robôs.

### **Necessidade de dicto/de re**

Necessidade de dicto é a espécie de necessidade que vincula palavras, por exemplo, que necessariamente todos os solteiros são não-casados, ou que necessariamente  $2 + 3 = 5$ . Necessidade de re é a espécie de necessidade de que se relaciona a coisas. Para entendê-la, precisamos usar a terminologia de mundos possíveis. Portanto, podemos dizer que há um mundo possível no qual Clinton nunca foi o presidente dos Estados Unidos, mas não há mundo possível no qual Clinton não fosse Clinton. Similarmente, ainda que haja um mundo possível no qual Clinton fosse careca, não há mundo possível no qual ele não fosse humano. Charles Dodgson pode não ter escrito *Alice in Wonderland* (*Alice no país das maravilhas*) em algum mundo possível, mas em todo mundo possível ele não pode ter sido outra pessoa além de Lewis Carroll.

### **Nomológico**

Do grego *nomos*, isto é, lei. Portanto, nomológico é “similar à lei”.

### **Normativo**

Governado por regras ou padrões.

### **Ocasionalismo**

A doutrina da relação mente/corpo associada ao filósofo francês Nicholas Malebranche (1638-1715). Malebranche sustentou que a mente, concebida como uma alma não-física,

não tem comércio causal com o corpo material, mas que, quando estados mentais e físicos ocorrem, Deus está presente em cada ocasião – por isso o título da doutrina – para garantir que as espécies corretas de correspondência entre o mental e o físico ocorram, de tal forma que é como se elas afetassem umas às outras. Malebranche adotou a doutrina para contornar o problema cartesiano de como coisas físicas e não-físicas poderiam, possivelmente, ter qualquer efeito uma sobre a outra.

### **Ocorrente**

Estados mentais ocorrentes são aqueles que estão atualmente presentes na consciência – por exemplo, a dor em meus ombros que estou experienciando neste momento. Estados mentais ocorrentes contrastam com estados **disposicionais** que se manifestam quando certas condições são preenchidas. Minhas crenças são disposicionais no sentido de que não estou atualmente consciente delas, mas posso chamá-las à mente se e quando a ocasião exigir, e nesse caso elas se tornam ocorrentes.

### **Ontologia/ontológico**

O braço da metafísica que se ocupa com a natureza do que existe.

### **Paralelismo psicofísico**

Esta teoria, inventada por Leibniz, foi elaborada, como o **ocasionalismo** de Malebranche, para escapar do problema posto pelo dualismo **cartesiano**, a saber, como é possível que uma alma não-física (espírito) possa afetar o corpo físico, e vice-versa. Eventos mentais e eventos físicos correm paralelamente um ao outro, e são sincronizados um com o outro. Isso porque Deus, que dá corda nos relógios do espírito e do corpo, sincroniza as espécies certas de eventos mentais com as espécies certas de eventos físicos, e então põe os relógios a funcionar. Ao contrário do Deus de Malebranche, o de Leibniz não precisa estar presente em cada uma e todas as ocasiões para assegurar a ocorrência das correspondências mentais e físicas apropriadas.

### **Platão/platônico**

Platão nasceu em Atenas aproximadamente em 428 a.C. e é mais renomado por seus

escritos sobre ética e sua teoria das Formas. De acordo com essa teoria, o mundo real, o que pode ser chamado de céu platônico, é um mundo de características universais – os universais – que são perfeitas, imutáveis e eternas, conhecidas apenas por meio da familiaridade intelectual após o treinamento em matemática, geometria e filosofia. O mundo que vemos à nossa volta é imperfeito, o mundo imperceptível e mutável de particulares percebidos pelos sentidos. Particulares são feitos para serem as espécies de coisas que são por compartilharem da natureza das Formas, ou por serem cópias imperfeitas ou reflexões delas. Conhecimento é familiaridade com as Formas não-sensórias, crença é familiaridade com particulares sensórios. Platão também aderiu ao **dualismo**, e pensou que a natureza da alma fosse similar àquela das Formas.

### **Positivismo lógico**

O positivismo lógico foi associado aos membros do Círculo de Viena, que prosperou entre 1924 e 1936, aproximadamente. Os positivistas lógicos estavam muito preocupados com a unidade da ciência e com o método científico. Um objetivo central foi formular um critério que faria a demarcação entre as sentenças significativas e aquelas desprovidas de sentido, e isso levou à formulação do **princípio da verificação**, de acordo com o qual uma sentença é significativa se, e somente se, ela puder ser verificada em princípio, tendo por base a experiência. Posteriormente, sob a pressão dos críticos, o princípio da verificação transmutou-se no princípio da falsificação, de acordo com o qual uma sentença é destituída de sentido literal, a menos que ela possa em princípio ser falsificada empiricamente.

### **Premissa**

Uma premissa é uma sentença que sustenta uma conclusão e que, supostamente, conduz a ela. Em **argumentos dedutivos** válidos, as premissas do **argumento** implicam logicamente a conclusão.

### **Psuche (plural: psuchai)**

Aristóteles fala de *psuche* como uma essência definidora de uma coisa viva, o conjunto de capacidades que torna a coisa em questão

a espécie de coisa que ela é. Plantas, animais não-humanos e seres humanos possuem *psuche*, mas, claramente, a capacidade de cada um difere da dos outros. Plantas possuem a capacidade de obter seus nutrientes e de crescer, mas não a de perceber ou se mover. Animais podem nutrir-se a si mesmos e crescer, e também podem perceber e se mover. Seres humanos, que também podem fazer todas as coisas que plantas e animais podem, também têm a capacidade de usar a razão, pois têm a capacidade de adquirir e usar a linguagem.

### **Qualia**

Qualia compreendem os sentimentos ou aspectos fenomenológicos da experiência, tais como a sensação de coceira produzida pelo colete, a incômoda dor de câimbra ou o verde vivo e dourado de um campo de arroz novo. Os qualia, supõem-se, são subjetivos e privativos à pessoa individual e existem além dos eventos públicos, físicos que ocorrem no cérebro e no sistema nervoso central.

### **Quantificadores existencial e universal**

Há dois quantificadores na lógica, o quantificador existencial ( $\exists x$ ), que diz "há ao menos uma coisa que é x...", e o quantificador universal estrito ( $\forall x$ ), que diz "para todas as coisas que são x...". Esses quantificadores são definíveis em termos um do outro e ocorrem no cálculo de predicados. Por exemplo, ( $\exists x$ ) ( $Fx \wedge Gx$ ) afirma que há ao menos uma coisa x tal que tenha a propriedade F e a propriedade G. Seja F = "é um homem" e G = "é careca". Então essa sentença diz que há ao menos um homem que é careca, isto é, alguns homens são carecas.

### **Redução analítica**

A ideia básica por trás da redução analítica é que falar sobre um certo tipo de coisas pode ser expresso, sem perda de sentido, em falar sobre outra classe de coisas de uma maneira que deixe claro o que está sendo assertado. Por exemplo, tem sido proposto que falar sobre objetos físicos pode ser traduzido, sem perda de sentido, em falar sobre experiências reais e possíveis que alguém possa ter. Similarmente, o behaviorismo analítico sustenta que

falar sobre a mente é traduzível, sem perda de sentido, em falar sobre comportamento público real e potencial. Veja também **redução ontológica**.

### **Redução ontológica**

Uma redução ontológica contrasta com uma **redução analítica**. Uma redução analítica sustenta que falar sobre uma espécie de coisas pode ser realizado com palavras diferentes sem perda de sentido. Consequentemente, falar sobre triângulos pode ser expresso, sem perda de sentido, em falar sobre figuras com três ângulos. As duas espécies de falar têm sentidos equivalentes. Uma redução ontológica nunca sustenta que descrições de um fenômeno e a análise dele são equivalentes quanto ao sentido, mas apenas que os termos que figuram na análise constituem todos os fatos sobre o fenômeno que está sendo reduzido ontologicamente. Consequentemente, falar sobre o calor não é analiticamente redutível a falar sobre o comportamento das moléculas – as duas espécies de falar têm sentidos diferentes, irredutíveis, no entanto, não há fatos sobre o calor no macroplano que estejam além do comportamento das moléculas no microplano.

### **Reificar/reificação**

Reificar um item é tratá-lo como uma coisa. Isso acontece comumente por causa do equívoco de pensar que todos os substantivos adquirem seus significados pelo ato de nomear coisas, mas muito frequentemente isso é um equívoco. A palavra “dor” é um substantivo, mas a dor é uma coisa? Antes, deveríamos pensar nisso como algum estado ou condição de uma pessoa. Similarmente, o sentido de uma palavra é um objeto, ou é melhor pensar nela, como Wittgenstein sugere, como o uso de uma palavra? O substantivo “corpo” nomeia uma coisa, mas há necessidade de que isto seja verdade quanto ao substantivo “espírito”, contrariamente ao que sustentam os **dualistas** de substância?

### **Res cogitans**

Uma coisa pensante, do latim *res*, “uma coisa” e *cogitans*, “pensante”. Era assim que Descartes caracterizava o espírito ou a alma.

Ele é uma coisa cuja essência em sua totalidade é pensar. Por pensamento ou, em francês, *penser*, Descartes quer dizer consciência em geral, e não uma mera inteleção ou raciocínio.

### **Res extensa**

Uma coisa extensa, isto é, física, corpos que ocupam espaço. Descartes dividiu todo o universo em duas esferas mutuamente exclusivas: **res cogitans** e **res extensa**.

### **Semântica e sintaxe**

Semântica é o estudo do significado de frases e de como isto é possível. Sintaxe é o estudo das estruturas da linguagem, por meio das quais as frases são produzidas.

### **Sentença de Ramsey**

Inventada pelo filósofo de Cambridge Frank Plumpton Ramsey (1903-1930), o objetivo de uma sentença de Ramsey é eliminar a circularidade na análise dos fenômenos. Por exemplo, para se eximir de definir um item como “carga positiva” em termos de seu oposto, “carga negativa”, e vice-versa, a saída do círculo é falar sobre ambas as cargas de modo mais neutro e sem compromisso, dizendo que há algo que possui a propriedade F e há algo que possui a propriedade G; coisas com F atraem aquelas com G e repelem coisas com F, enquanto coisas com G atraem coisas com F e repelem coisas com G; e G pode ser induzido em um pente de borracha, esfregando-o com lã.

### **Silogismo**

Um silogismo é uma forma de inferência dedutiva na qual uma conclusão é extraída de duas premissas. Por exemplo, “Todo peixe vive na água”, “Todas as coisas que vivem na água são de sangue-frio”, consequentemente, “todos os peixes são de sangue-frio”. Cada premissa contém uma expressão que ocorre na conclusão, e uma expressão, chamada de “termo médio”, que não ocorre. Nesse caso, o termo médio é “vive na água”. Há quatro formas básicas de afirmações que podem figurar em silogismos: Essas são: “Todos os A são B”, “Nenhum A é B”, “Alguns A são B” e “Alguns A não são B”. Há 256 formas nas quais essas afirmações podem ser combinadas, mas

apenas 14 dessas combinações representam argumentos válidos, isto é, argumentos nos quais a conclusão se segue logicamente das premissas.

### **Sobredade**

Um termo informal para estados mentais que exibem **intencionalidade**, a característica de ser representativo de outros estados de coisas, ou ser relativo a eles, incluindo aqueles que não existem. Em outras palavras, estados mentais intencionais, tais como crenças, sonhos, desejos e pensamentos, representam o mundo como sendo de certa maneira, e então possuem conteúdo representacional. Por exemplo, minha crença de que há fadas no fundo de meu jardim representa meu jardim como possuindo fadas nele. Mesmo que nada na realidade corresponda a essa crença, a crença é relativa a isso. Quando o mundo é como a crença o representa como sendo, a crença é verdadeira.

### **Solipsismo**

A visão de que apenas o próprio eu e suas experiências existem. Uma versão fraca acrescentaria “tanto quanto o eu possa dizer”.

### **Substância/substância lógica**

Uma substância é uma coisa que pode existir por si mesma, com independência lógica de tudo o mais, diferentemente das propriedades que têm de ser características de uma ou outra substância. Contudo, assim como as propriedades precisam pertencer às substâncias, pareceria igualmente que não podem existir substâncias sem características ou propriedades. Substâncias e propriedades são feitas uma para a outra. É um casamento feito no céu metafísico que nenhum filósofo pode separar. Alguns filósofos, por exemplo, Espinosa (1632-1677), têm sustentado que, estritamente falando, Deus é a única substância genuína, porque Deus, ao contrário de tudo mais, não depende para sua existência de nada externo a ele.

### **Superveniente/subveniente**

Um fenômeno superveniente surge de uma base subveniente e dela depende para sua

existência. Além disso, não podem ocorrer mudanças no fenômeno superveniente com uma correspondente mudança na base subveniente, mas o oposto não é válido devido à possibilidade de uma realização múltipla do fenômeno superveniente por uma variedade de bases subvenientes. A natureza exata da relação superveniente é controversa.

### **Tautologia**

Uma tautologia é uma espécie de verdade necessária porque ela meramente repete a si mesma. Por exemplo, o verso “Uma rosa é uma rosa é uma rosa” (da poetisa Gertrude Stein) é uma tautologia.

### **Verificacionismo/princípio da verificação**

O verificacionismo é uma doutrina central do **positivismo lógico** que sustenta que, a menos que uma sentença possa ser verificada empiricamente, isto é, estabelecida como verdadeira em princípio, tendo por base a experiência, ela é destituída de sentido. (Isso não se aplica a verdades e falsidades necessárias.) O problema é que existem todas as espécies de afirmações que não podem ser verificadas e ainda assim fazem sentido – por exemplo, uma cidade nunca será construída neste local (um exemplo dado pelo filósofo americano Hilary Putnam). Além disso, como alguém pretenderia verificar o próprio princípio da verificação – por observação ou experimentação? O principal erro produzido pelo princípio é que, enquanto deve ser possível especificar que estado de coisas, caso tenha ocorrido, faria com que uma sentença fosse verdadeira para que ela seja dotada de sentido, é uma questão totalmente diferente ser realmente capaz de descobrir, mesmo em princípio, se aquele estado de coisas ocorre. Podem existir muitos tipos de afirmações, cuja verdade nós nunca seremos capazes de descobrir, mesmo em princípio, mas isso não faz com que elas sejam desprovidas de sentido.

## REFERÊNCIAS

---

- Aristotle, *De Anima*, Books II and III, translated with an introduction and notes by D.W. Hamlyn (Oxford: Oxford University Press, 1968).
- Armstrong, D. M., *A Materialist Theory of the Mind* (London: Routledge and Kegan Paul, 1968).
- Butler, Joseph, "Of Personal Identity", in J.H. Bernard (ed.), *The Works of Bishop Butler*, vol. II (London, 1900).
- Chalmers, David and Searle, John, "Consciousness and the Philosophers: An Exchange", *New York Review of Books* (15 May 1997), 61.
- Churchland, Patricia, *Neurophilosophy* (Cambridge, MA: MIT Press, 1986). Churchland, Paul, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979).
- Cummins, Robert, *Meaning and Mental Representation* (Cambridge, MA: MIT Press, 1995).
- Davidson, Donald, "Thought and Talk", in Samuel Guttenplan (ed.), *Mind and Language: Wolfson College Lectures 1974* (Oxford: Clarendon Press, 1975). Dennett, Daniel, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987). Dennett, Daniel, *Consciousness Explained* (Boston, MA: Little, Brown, 1991). Dennett, Daniel, *Kinds of Minds: Toward an Understanding of Consciousness* (New York: Basic Books, 1996).
- Dennett, Daniel and Hofstadter, Douglas (eds), *The Mind's I* (New York: Basic Books, 1981).
- Descartes, René, *Meditations on First Philosophy*, trans. John Cottingham (Cambridge: Cambridge University Press, 1986 [1641]).
- Evnime, Simon, *Donald Davidson* (Cambridge: Polity, 1991).
- Fodor, Jerry, *The Language of Thought* (New York: T. Y. Crowell, 1975). Fodor, Jerry, *Psychosemantics* (Cambridge, MA: MIT Press, 1988).
- Haugeland, John, *Artificial Intelligence: The Very Idea* (Cambridge, MA: MIT Press, 1985).
- Heil, John, *The Nature of True Minds* (Cambridge: Cambridge University Press, 1992).
- Hume, David, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1978 [1739]).
- Jackson, Frank, "Mental Causation", *Mind* 105 (1996), 377-41.
- Kant, Immanuel, *The Critique of Pure Reason*, trans. Norman Kemp Smith (London: Macmillan, 1964 [1787]).
- Kim, Jaegwon, *Supervenience and Mind: Selected Philosophical Essays* (Cambridge: Cambridge University Press, 1993).
- Kirk, Robert, "Zombies vs. Materialists", *Proceedings of the Aristotelian Society*, supplementary vol. 48 (1974), 135-52.
- Kirk, Robert, *Raw Feeling* (Oxford: Clarendon Press, 1996).
- Kripke, Saul, *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1980).
- La Mettrie, Julien Offray de, *Man a Machine*, trans. R. Watson and M. Rybalka (Indianapolis: Hackett Publishing Co, 1994 [1747/8]).
- Leibniz, Gottfried Wilhelm, *Monadology* (1787), repr. in Mary Morris and G.H.R. Parkinson (trans.), *Philosophical Writings* (London, 1973), 94.
- Locke, John, *An Essay Concerning Human Understanding*, ed. P.H. Nidditch (Oxford: Clarendon Press, 1978 [1690]).
- Lockwood, Michael, *Mind, Brain, and Quantum* (Oxford: Basil Blackwell, 1989). Lockwood, Michael, "The Grain Problem", in H. Robinson (ed.), *Objections to Physicalism* (Oxford: Clarendon Press, 1993), 271-91.

- Lowe, F. J., *Locke on Human Understanding* (London: Routledge, 1995).
- Lowe, F. J., *Subjects of Experience* (Cambridge: Cambridge University Press, 1996). Lycan, W. C., *Consciousness* (Cambridge, MA: MIT Press, 1987).
- Mackie, John, *Problems from Locke* (Oxford: Oxford University Press, 1976), esp. ch. 6.
- Nagel, Thomas, "What is it Like to be a Bat?", *Philosophical Review* 83 (1974), 435-50.
- Nagel, Thomas, *Mortal Questions* (Cambridge: Cambridge University Press, 1979). Nozick, Robert, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981), esp. ch. 1.
- Parfit, Derek, "Personal Identity", *The Philosophical Review* 80 (1971), 3-27.
- Perner, Josef, *Understanding the Representational Mind* (Cambridge, MA: MIT Press, 1991).
- Perry, John, "Can the Self Divide?", *Journal of Philosophy* 69 (1972), 463-88.
- Perry, John (ed.), *Personal Identity* (Berkeley: University of California Press, 1975). Place, U. T., "Is Consciousness A Brain Process?", *The British Journal of Psychology* 47 (1956), 44-50.
- Poland, Jeffrey, *Physicalism: The Philosophical Foundations* (Oxford: Clarendon Press, 1994).
- Putnam, Hilary, *Mind, Language, and Reality: Philosophical Papers*, vol. 2 (Cambridge: Cambridge University Press, 1975).
- Putnam, Hilary, *Reason, Truth, and History* (Cambridge: Cambridge University Press, 1981).
- Quinton, Anthony, "The Soul", *Journal of Philosophy* 59 (1962), 393-403.
- Rorty, Amélie (ed.), *The Identities of Persons* (Berkeley: University of California Press, 1976).
- Rosenthal, David (ed.), *The Nature of Mind* (New York: Oxford University Press, 1991).
- Ryle, Gilbert, *The Concept of Mind* (London: Hutchinson, 1949).
- Searle, John, *Minds, Brains, and Science. The 1984 Reith Lectures* (London: Penguin, 1989).
- Searle, John, *The Rediscovery of the Mind* (Cambridge, MA: MIT Press, 1992).
- Searle, John, "Consciousness and the Philosophers", *New York Review of Books* (6 March 1997), 43-50.
- Shoemaker, Sydney, "Causality and Properties", in Peter van Inwagen (ed.), *Time and Cause* (Dordrecht: Reidel Publishing Co, 1980), 109-35.
- Skinner, B. F., *Science and Human Behavior* (New York: Macmillan, 1953).
- Sterelny, Kim, *The Representational Theory of Mind: An Introduction* (Oxford: Blackwell Publishers, 1990).
- Stich, Stephen, *From Folk Psychology to Cognitive Science: The Case Against Belief* (Cambridge, MA: MIT Press, 1983).
- Strawson, P. F., *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen, 1959).
- Stroud, Barry, *Hume* (London: Routledge, 1977).
- Swinburne, Richard, "Personal Identity", *Proceedings of the Aristotelian Society* 74 (1973-4), 231-48.
- Turing, Alan, "Computing Machinery and Intelligence", *Mind* 59 (1950), 434-60. Watson, J. B., 'Psychology as the Behaviorist Views It', *Psychological Review* 20 (1913), 158-77.
- Wiggins, David, *Identity and Spatio-temporal Continuity* (Oxford: Oxford University Press, 1967).
- Wiggins, David, *Sameness and Substance* (Oxford: Oxford University Press, 1980). Williams, Bernard, *Problems of the Self* (Cambridge: Cambridge University Press, 1973).
- Wittgenstein, Ludwig, *Tractatus Logico-Philosophicus*, trans. D. F. Pears and B. R. McGuinness (London: Routledge and Kegan Paul, 1961 [1922]).
- Wittgenstein, Ludwig, *The Blue and Brown Books* (Oxford: Oxford University Press, 1958).
- Wittgenstein, Ludwig, *Philosophical Investigations*, trans. G. E. M. Anscombe (Oxford: Basil Blackwell, 1968 [1953]).
- Wittgenstein, Ludwig, *On Certainty*, eds G. E. M. Anscombe and C. H. von Wright, trans. Denis Paul and G. E. M. Anscombe (Oxford: Basil Blackwell, 1969).

# ÍNDICE

## A

*À la recherche du temps perdu* (Proust) 29-30  
Acesso privilegiado 27-28, 33-34  
Behaviorismo analítico 107-108  
Teoria da identidade 86-87  
**Ações**  
Intencionalidade 20-22  
Racionalizações 186-188  
Teoria causal das ações intencionais de Davidson 188-193  
"Ações, razões e causas" (Davidson) 183-184, 187-188  
**Agentiva**  
Descrição agentiva 287  
Descrição comportamental 110-111  
**Alma, a**  
Argumento favorável de Platão 41-45  
Comércio causal com o corpo 159-160  
Dois testes cartesianos a favor 48-50  
Dualismo cartesiano 14, 47-48  
Dualismo de substância 39-41  
Individuação e identificação ao longo do tempo 159-160  
O todo e não partes 280-281  
Um ou muitos? 66-68  
*Analysandum e analysans* 108-109  
Animais, estados mentais de 106-107  
Anscombe, Elizabeth  
Crença e emoção 20-22  
*An Anthropologist on Mars* (Sacks) 37n, 157n  
Antirrealismo 238-240  
Wittgenstein e estados mentais  
**Aparência**  
Conteúdo intencional de percepções 22-23  
**Argumento**  
Dedutivo 42-43  
Silogismos 43  
Ver também: argumento dedutivo;

## Argumento indutivo

Argumento dedutivo 222-223, 238-239, 285  
Ver também argumento indutivo  
Argumento indutivo 42-43, 238-239, 246n, 285, 289-290  
Argumento da analogia 223  
Hume justifica 241  
Outras mentes 223-225  
Wittgenstein ataca argumento da analogia 224-227  
Ver também: argumento dedutivo  
Argumento do quarto chinês 150-154  
Aristóteles 201-218  
Animais 205-206  
Capacidades 204-212  
Consciência 211-213  
*De Anima* 204-205, 212-213  
Defines substancia 40-41  
Doutrina do intelecto ativo 212-214  
E funcionalismo 207-208  
Falácia mereológica 214-218  
*Geração de animais* 213  
Interação mente/corpo 214-215  
Línguagem 205-207  
Materia e forma 202-204  
Natureza política do homem 34-35  
*Psiche* 203-208, 214-218  
Raiva 210-211  
Sensíveis especiais 205-206  
Silogismos 43  
Armstrong, David (1926-) 72-73  
A resposta de Searle a 169-170  
Arnauld, Antoine (1612-94)  
Objecção a Descartes 60-61  
**Assertivas de verdade**  
Argumentos a partir de premissas 42-43  
Ryle e verificionismo 116-118  
Teoria da identidade 96-101  
Ayer, A.J. (1910-89)  
Sobre a dúvida cartesiana 54-55  
*The Concept of a Person* 243-244

## B

Behaviorismo 219n  
Comparado ao funcionalismo 143-144  
Estados mentais 31-32  
Metodológico 106-108  
Ver também: behaviorismo analítico  
**Behaviorismo analítico**  
Agentiva 110-111  
Behaviorismo "brando" de Ryle 116-119  
Behaviorismo "duro" de Hampel 109-113  
Cinco elementos de análise 111-112  
Circularidade e regressão infinita 114-117  
Comparada à teoria da identidade  
Comparada ao dualismo 105-106  
Consciência 159-160  
Contrastada com o behaviorismo metodológico 106-108  
Crenças abstratas complexas 107-108  
Desejos não causas de ações 184-187  
E Wittgenstein 238-240  
Estados mentais disposicionais 119-120  
Negação de estados mentais subjetivos 121-123  
Negação do conhecimento de primeira pessoa 122-123  
Padrões de comportamento 112-115  
Pontos fortes da teoria 105-107  
Reducionismo 221-223  
Tragos de 107-110  
Zumbis e superespartanos 122-125  
**Behaviorismo de estímulos e respostas** 107-108  
Biopsicologia 202  
Block, Ned 139-141  
Argumento da mente chinesa 144-146  
Argumento das cabeças de homúnculos 144-146

- Blue Book, The* (Wittgenstein) 229-230, 233-234
- Boswell, James 118-119
- Brentano, Franz (1838-1917)  
Conceito de intencionalidade 22-25
- Broca, Pierre 78-79, 208-209
- Burge, Tyler  
Epifenomenalismo 183-184
- Burwood, S. 163-164
- Butler, Bishop Joseph  
Consciência 173  
Eu não-reducionista 279-280  
Objecção à identidade pessoal de Locke 267-268  
"Tudo é o que é" 249-250
- C**
- Cartesianismo. Ver Descartes, René
- Causação psicofísica 286  
"Actions, Reasons and Causes" de Davidson 183-184, 186-188  
Argumento de Malcolm 184-187  
Monismo anômalo 192-194  
Problemas de epifenomenalismo 183-184  
Rejeição da sobredeterminação 197-199  
Teoria causal de ações intencionais de Davidson 188-193  
Variações de relações causais 181-182
- Causalidade 283n, 286  
Análise de Hume de 108-109, 186-186, 126n  
Behaviorismo e "sintomas" 120  
Covariância 89-91  
Desejos causando ação 183-185  
Efeitos 191-192  
Funcionalismo 143-144  
Monismo anômalo 192-194, 196-199  
Psicofísico 49-53, 286  
Superveniência 194-197  
Teoria causal da ação intencional de Davidson 188-193  
Teoria humeana do feixe do eu 271-272  
Variedades de estados físicos e mentais 182-183  
Ver também causação psicofísica
- Cérebro, O  
Acesso privilegiado 221-222  
Aspecto espacial 81-84  
Atividade neural 61-62  
Células individuais de 155-156, 277-278  
Cisão 273-277, 279-281  
Como um objeto físico 13-14  
Dualismo 14  
Dualismo de substância 279-281
- Escaneamentos PET e MRI 78-79  
Intencionalidade 35-36  
Mau funcionamento de 65-66  
Mecanismo de transferência de estado cerebral de Parfit 273-277  
Naturalismo biológico 164-166  
Percepção de cores 157n  
Perspectiva pessoal de 25-27  
Preservador de identidade psicológica 277-279  
Processos de acordo com o dualismo de propriedade 173-178  
Psicofuncionalismo 131-136  
Teoria da identidade 14-15  
Transplantes e identidade física 234-236
- Ceticismo  
Epifenomenalismo 183-184
- Chalmers, David (1966- )  
Consciência 174-178  
*The Conscious Mind* 104n
- Chomsky, Noam (1928-)  
Dispositivo para a aquisição da linguagem (DAL) 287  
O cérebro e linguagem 48-50
- Churchland, Patricia (1943- ) 101-102
- Churchland, Paul (1942- ) 101-102  
Sobre o argumento de Jackson 149-151
- Ciência  
Metodologia e os positivistas lógicos 109-110
- Circularidade 114-117  
Funcionalismo 142-143
- Círculo de Viena 292  
Positivismo lógico 109-110
- Cognições 16-17
- Comportamento  
E estado mental 221-223  
Substituição de termos mentais 239-241
- Computadores  
Comparados a mentes 136-142  
Intencionalidade 290  
Quarto chinês de Searle 151-154  
Ver também inteligência artificial
- Conceitos, formação humana de 173-174
- Concept of a Person, The* (Ayer) 243-244
- Concept of Mind, The* (Ryle) 45-46, 116-117
- Conhecimento  
De próprios estados mentais 27-29  
Estados conativos 18-19
- Conrad, Joseph  
*Lord Jim* 29-31, 186-187
- Consciência  
Células cerebrais individuais 155-156
- Conceito de pessoalidade de Ayer 243-244  
Dúvida cartesiana 57-58  
Intencionalidade 33-35  
Monismo anômalo 195-196  
Naturalismo biológico 167-168  
Natureza da 172  
O não-consciente 47-48  
Omissões de teorias 159-160  
Processos cerebrais físicos 173-178, 211-213  
Responsável pela coerência 35-36  
Searle sobre 169-171  
Wilkes sobre 211-212  
Ver também: continuidade psicológica
- Conscious Mind, The* (Chalmers) 104n
- Continuidade psicológica  
Argumento favorável de Unger 276-279
- Butler objeta Locke 267-268
- Deus 275-276
- Dispositivo de transferência de estado cerebral de Parfit 273-277
- Não-reducionismo 278-283
- Objecção da circularidade 267-271
- Parfit modifica Locke 266-271
- Reid objeta Locke 265-266
- Sem continuidade física 273-277
- Teoria de Locke 264-266
- Cor  
Impedimento cerebral 157-158n  
Locke sobre 157n  
Movimentos corporais "sem cor" 110-111, 114-115
- Qualia invertido 146
- Significados privados 225-226
- Teoria da identidade espécime-espécime 94-95
- "What Mary Didn't Know" de Jackson 147-151
- Wittgenstein sobre definição ostensiva 229-232
- Cottingham, John  
Pensamento e atividade cerebral 62-64
- Crane, Tim  
*The Mechanical Mind* 137-140
- Crerias  
Behaviorismo analítico 113-115  
Consciência e intencionalidade 33-35  
Conteúdo de emoções 20-22
- Disposições 119-120
- Estados conativos 18-19
- Racionalidade e teoria da identidade 97-99
- Segunda ordem 30-31
- Crick, Francis 76-77
- Critique of Pure Reason, The* (Kant) 66-67, 226-227

- Cummings, Robert  
*Meaning and Mental Representation*  
 217-218
- D**
- Davidson, Donald (1917-) 14-15  
 "Actions, Reasons and Causes" 173,  
 186-188  
 Doutrina da superveniência 160-163  
 Monismo anômalo 181-182, 192-194  
 Objeções de Kim ao monismo  
 anômalo 194-195  
 Racionalidade e teoria da identidade  
 93-94  
 Teoria causal da ação intencional  
 188-193
- Dennett, Daniel (1942-)  
 Homúnculos e funcionalismo 154-156  
 Identidade aparente 56-57
- Descartes, René (1596-1650)  
 Argumento da divisibilidade 53, 63-  
 65  
 Argumento da dúvida 53-58  
 Argumento da percepção clara e  
 distinta 53, 57-64  
*Cogito ergo sum/ penso, logo existo*  
 53-55  
 Coisas espaciais e mentais 81-83  
 Conjecturas sobre a glândula pineal  
 50-51, 66  
 Dualismo 14, 214-215  
 Dualismo de propriedade 171-172  
 Dualismo mente/corpo 44-48  
 Estar consciente de 204-205, 276-277  
 Eu- e outras atribuições de estados  
 mentais 243-244  
 Independência do corpo e alma 164-  
 165, 203-205, 214  
 Interpretações de 70-71n  
 Materialismo 278  
*Meditations* 46-47  
 Mentes de animais 133, 224-225  
 O "fantasma na máquina" 116-118  
 Pessoas desencarnadas 124-125  
 Privacidade e acesso privilegiado 27-  
 28, 31-32, 224-226, 227-229  
 Relação de mente com corpo 49-53  
*Res cogitans* 293  
 Significados de imagens 232-234  
 Teste para a presença da alma 48-50,  
 208-209  
*The Discourse on Method* 46-49  
*The Principles of Philosophy* 47-48  
 Uma alma 66-68
- Desejos de segunda ordem 134  
 Desenfreados 134  
 Dickens, Charles  
*Hard Times* 37n  
*Discourse of the Method, The* (Descartes)
- 46-49  
 Disposição  
 Behaviorismo de Ryle 119-120
- Dor 37m  
 Abordagem criterial 236-238  
 Autojulgamento de 28-30  
 Behaviorismo analítico 112-114, 122-  
 124  
 Condição individual de 83-84  
 Crítica de Kripke à teoria da  
 identidade 98-101  
 De membro fantasma 17-19  
 Estar consciente de 126n  
 Experiência privada 32-33  
 Fingimento 122-125  
 Funcionalismo 143-145  
 Membros fantasmas 17-19  
 Monismo anômalo 195-197  
 Objeções de Wittgenstein ao  
 argumento indutivo 224-227  
 Teoria espécime-espécime 80-81  
 Teoria tipo-tipo 77  
 Wittgenstein e linguagem de 239-241
- Doutrina da superveniência 179n, 294  
 Aristóteles e 215-216  
 Caverna de Platão 162-163  
 Co-variação 161-163  
 Dependência 161-163  
 Irreduzibilidade 161  
 Monismo anômalo de Davidson 193-  
 194  
 Moralidade e estética 163-165  
 Natureza da relação 162-164  
 Searle 169-172, 194-196  
 Teoria da identidade espécime-  
 espécime 94-95  
 Traços de 160-163
- Dualismo de propriedades  
 Comparado à substância 40-42  
 E epifenomenismo 171-172  
 Monismo não-reduutivo 160  
 Natureza de propriedades mentais  
 não-físicas 171-173  
 Processos cerebrais 171-178  
 Searle e superveniência 169-172
- Dualismo  
 Argumento básico 14  
 Argumento da percepção definida  
 57-64  
 Argumento de Ryle 45-46, 117-118  
 Cartesiano 46-48, 286  
 Contando almas 66-68  
 Definido 288  
 Falácia do homúnculo 65-66  
 Função cerebral 65-66  
 Identidade pessoal não-reduativa 278-  
 283  
 Mentalidade tudo ou nada 64-65  
 Platônico 40-42, 43-47, 292  
 Propriedade 40-42, 160, 170-178
- Substância 39-42, 159-160, 209-210,  
 279-283
- Dúvida  
 Cartesiana 53-58  
 Estados mentais versus  
 comportamento 241
- E**
- Eccles, Sir John 214-215  
 Elliot, T.S.  
 "Morning at the Window" 40-41  
 Emoções 16-17  
 Conteúdo 20-22  
 Desejos ignorados no  
 epifenomenismo 183-184  
 Estados conativos 19-21  
 Sensações associadas ao 19-21  
*Enquiry Concerning Human  
 Understanding* (Hume) 185-186, 271-  
 272
- Epicuro (324-270 a.C.)  
 Materialismo 72-73
- Epifenomenismo 178-179  
 Definido 288  
 E dualismo de propriedade 171-172  
 Problemas do 183-184  
 Rejeitado por Davidson 192-194
- Epistemologia 288  
 Definida 220-221  
 Existência de estados mentais 15-16  
 Identidade pessoal 249
- Escola pitagórica 60-61
- Espaço  
 Aspectos da teoria da identidade 81-  
 84
- Estados conativos 16-17  
 Estar consciente sem sensações 18-20
- Estados de quase-percepções 16-17
- Estados físicos  
 E a alma 43-46  
 Espacialidade 36  
 Existência sem termos consciência de  
 35-36  
 Fenômenos 289  
 Intencionalidade e eventos cerebrais  
 35-36
- Leis causais dos 188-193  
 Perspectivas de terceira pessoa 35-36  
 Pessoas desencarnadas 124-125  
 Relação a estados mentais 13-14  
 Sensações 16-19  
*Ver também:* teoria da identidade
- Estados mentais  
 Atribuição-eu-e-outro de Strawson  
 241-244  
 Behaviorismo analítico 106-107  
 Características de 15-17

- E comportamento 221-223  
 Funcionalismo 130-132, 142-144  
 Holístico versus atomístico 285-286  
 Intencional 87-93  
 Não-intencional 288  
 Natureza do 12-14  
 Negação behaviorista da subjetividade interna 120-121  
 Ocorrente 291-292  
 Perspectivas de primeira e de terceira pessoas 25-28, 30-33, 35-36, 122-123, 250-252  
 Relação a estados físicos 13-14  
 Relação cartesiana da mente ao corpo 49-53  
 Variedades de causação 182-183  
 Verificabilidade 109-111  
 Vocabulário para 100-102
- Estética  
 Doutrina da superveniência 164-165
- Eventos. Ver: causalidade
- Evolução  
 Teoria da identidade 84-86
- Experiência  
 Privacidade da 25-28  
 Problema mente/corpo 13-14  
 Teoria humeana do feixe do eu 270-274
- F**  
 "Fantasma na máquina"  
 Crítica de Ryle 117-119
- Fenomenalismo 289  
*Analysandum e analysans* 108-109
- Fenomenologia  
 De sensações 19-20
- Fenômenos  
 Doutrina da superveniência 161-163  
 Físicos 289
- Fenômenos mentais 12-13  
 Como fenômenos físicos 73-74  
 Definidos 289  
 Intencionalidade 22-25
- Fisher, David 214
- Fiscalismo 289  
 Teoria da identidade 73-75
- Flanagan, Owen  
*Consciousness Reconsidered* 180n
- Flew, Antony 210-212, 275-276  
 Dualismo platônico 44-45
- Fodor, Jerry (1935- )  
 "Mentalés" 88-89, 216-217
- Formas  
 Teoria da identidade espécime-espécime 94-95
- Formas, platônicas
- Definidas 41-42  
 Foster, John  
 Behaviorismo analítico 113-116
- Frankfurt, Harry 134
- Frege, Gottlob (1848-1925)  
 Pensamento privado 226-227
- Funcionalismo 14-15, 289  
 Argumento da mente chinesa de Block 144-146  
 Argumento do quarto chinês de Searle 150-154  
 Comparado ao behaviorismo 143-144  
 Computacional 136-142, 286  
 E Aristóteles 207-208  
 E homônimos 144-146, 154-156  
 E materialismo 147  
 Estar consciente de 159-160  
 Funções definidoras 129-131  
 Liberalismo 144-145  
 Mente como software 62-64  
 Metafísico 130-133, 151-152  
 Pontos fortes da teoria 143-144  
 Psico 131-136  
 Qualia invertido 146-147  
 Reducionismo 142-144  
 "What Mary Didn't Know" de Jackson 147-151
- G**  
 Garratt, Brian 278-279  
 Gassendi, Pierre (1592-1655)  
 Função cerebral 65-66
- Gilbert, P. 163-164
- God  
 Argumento cartesiano da percepção distinta 57-59, 61-62  
 Causa de pensamento 53-54  
 Identidade psicológica 275-276  
 Papel na interação mente/corpo 52-53
- Graham, George  
*Philosophy of Mind: An Introduction* 243-244
- H**  
 Hacker, Peter 210-213, 237-238  
 Falácia mereológico 216-217  
 Teoria da conexão cerebral 214-215
- Hahn, Oswaldo (1927-2005)  
 Abordagem criterial 237-238  
 Sobre Wittgenstein 240
- Hard Times (Dickens) 37n
- Hare, Richard, M. (1919-2002)  
*The Language of Morals* 163-164
- Heil, John  
 Sobre estados mentais próprios 27-28
- The Nature of True Minds* 30-31
- Hempel, Carl (1905-97)  
 Behaviorismo "duro" 109-113  
 Cinco elementos de análise 111-112  
 E Ryle 118-119
- Hobbes, Thomas (1588-1679)  
 Crítica o argumento cartesiano da dúvida 57-58  
 Descrença em substâncias incorpóreas 73-74
- Holbach, Baron d' (1723-89)  
*The System of Nature* 73-74
- Homônimos  
 Falácia 65-66, 288-289  
 No funcionalismo 154-156  
 Principal argumento de Block 144-146
- Hopkins, James 32-33
- Hull, C.L. (1884-1952) 110-111
- Hume, David (1711-76) 232-233  
 Aquisição de conceito 91-92  
 Causalidade 108-109, 166-167, 177-178, 187-188  
 Conceitos e experiência 67-68  
 Conhecimento da própria mente 27-28  
*Enquiry Concerning Human Understanding* 185-187, 271-272
- Estar consciente e processos cerebrais 176-177
- Formação de conceitos 173-174
- Justificação da indução 241
- Relações causais 126n
- Segunda definição de causa 185-186
- Teoria do feixe do eu 270-274
- Identidade  
 Aspectual 56-57  
 Continuidade pessoal 15-16  
 Lei de Leibniz 56-57  
 Ver também: identidade pessoal; teoria da continuidade física; continuidade psicológica
- Identidade pessoal 247-248  
 Aspectos epistemológicos e metafísicos 248-249  
 Definida 289  
 Mackie sobre 284n  
 Memória 251-252  
 Não-reducionismo 278-283  
 Qualitativo versus numérico 248  
 Reducionismo 249-253  
 Teoria da continuidade física 252-262
- Identidade, teoria da 14-15  
 Argumentos contrários de Kripke 96-101, 104n  
 Desenvolvimento histórico 72-75

- Espécime-espécime 78-81, 93-97, 112-113  
 Estados mentais intencionais 87-93  
 Estar consciente de 14-15, 168-169  
 Evidência do escaneamento cerebral 80-81  
 Materialismo eliminativo 100-103  
 Perspectiva de Aristóteles 201-202  
 Pontos fortes de 80-82  
 Problema mental e espacial 81-84  
 Qualia, acesso privilegiado e irreduzibilidade 85-88  
 Racionalidade e normatividade 93-94  
 Redução analítica 74-77  
 Reducionismo 179n  
 Reducionismo e espécime-espécime 93-97  
 Simetria de assertivas 83-86  
 Tipo-tipo 76-79, 96-97
- Imagens  
 Interpretação e sintaxe 232-234
- Indicação confiável 90-93
- Individuals* (Strawson) 241-244
- Indivíduos  
 Perspectivas de primeira pessoa 25-28, 30-33
- Inteligência artificial  
 O argumento do quarto chinês de Searle 151-154  
 Teoria reducionista da identidade 95-96
- Inteligência artificial* (filme) 120-121
- Intencionalidade  
 Atitudes proposicionais 23-25  
 Autojulgamentos de conhecimento 29-30  
 Como estados físicos 35-36  
 Conceito revivido por Brentano 23-25  
 Conteúdo de 20-23  
 Direção de ajuste 20-21  
 E desejo 185-187  
 Estados metais e teoria da identidade 87-93  
 Estados metais não intencionais 288  
 Estar consciente 33-35  
 Holismo 33-34, 93-94, 289  
 Homúnculos e funcionalismo 155-156  
 Intrínseco e derivativo 155-156, 290  
 Na ausência de ter consciência de 35-36  
 Naturalizado 24-26  
 Percepções 21-23  
 Searle 169-171  
 Sobredade 294  
 Teoria causal da ação de Davidson 188-193  
 "Volume de experiência" de Wittgenstein 20-21
- Intencionality* (Searle) 167-168
- Jackson, Frank (1943- )  
 Argumento de Churchland 149-151  
 "Qualia epifenomenica" 147-148  
 "What Mary Didn't Know" 147-151
- James, William (1842-1910)  
 O cérebro e estar consciente de 155-156
- Jaquette, Dale  
 Intencionalidade 88-89
- Jogos e esportes  
 Behaviorismo analítico e movimentos esportivos 114-115  
 Definição ostensiva 231-233
- Johnson, Dr. Samuel (1709-84) 118-119
- Jornadas nas Estrelas, 120-121, 132-133
- Joyce, James (1882-1941)  
*Ulysses* 48-49
- K
- Kant, Immanuel (1724-1804)  
 Conceitos vazios 2226-227  
 Contando almas 74-77  
*The Critique of Pure Reason* 66-67, 226-227
- Kekulé von Stradonitz, August 82-84
- Kenny, Anthony (1931- )  
 Intencionalidade intrínseca 153-156  
 Sobre a "percepção distinta de Descartes" 57-58  
*The Metaphysics of Mind* 153-155
- Kim, Jaegwon (1934- )  
 Doutrina da superveniência 161-163  
 Objeções ao monismo anômalo 194-197
- Kripke, Saul (1940- )  
 Crítica da teoria da identidade 96-101  
*Naming and Necessity* 96-97  
 Pessoas desencarnadas 124-125  
 Teoria da identidade 179n
- L
- La Mettrie, Julien Offray de (1709-51)  
*Man as a Machine and Man as a Plant* 73-74
- Language of Morals, The* (Hare) 163-164
- Lei de Leibniz 55-57, 149-150, 195-196  
 Definida 290  
 Estados mentais não estados cerebrais 220-221  
 Peças compostas e totalidades 278
- Leibniz, Gottfried (1646-1716)  
 Continuidade física 263n  
 O cérebro como uma máquina 13-14, 136
- M
- McGinn, Colin (1950- )  
 Causalidade 283n  
 Estados mentais 32-33  
 Estar consciente 173-178  
 "Mercados de mistérios" 180n

Significados de termos mentais 31-32  
 Suficiência e causalidade 271-272  
*Teste de Turing* 120-121  
 Mackie, J.L. (1977-81)  
 Pensadores e pensamentos materialistas 281-282  
 Sobre a identidade pessoal 263n, 284n  
 Malcolm, Norman (1911-90)  
 Desejos não causas de ações 183-187  
 Sobre o materialismo eliminativo 101-102  
 Sobre sonhos 126n  
 Malebranche, Nicholas (1638-1715) 291  
 Ocasionalismo 52-53  
*Man a Machine* (La Mettrie) 73-74  
*Man a Plant* (La Mettrie) 73-74  
*Man Who Mistook his Wife for a Hat, The* (Sacks) 126n  
 Marx, Karl (1818-83)  
 Materialismo cartesiano 278  
 Materialismo 290  
 A rejeição de Searle do materialismo reduutivo 169-172  
 Cartesiano 278  
 Eliminativo 100-103  
 Epicuro 72-74  
 Funcionalismo 14-15, 147-148  
 Mente como um software 62-64  
 Pensadores e pensamentos 281-282  
*Ver também:* teoria da identidade  
 Materialismo eliminativo 100-103, 288  
 Estar consciente 159-160  
 Popper sobre 104n  
*Mechanical Mind, The* (Crane) 137-140  
*Meditations* (Descartes) 46-47  
 Memória  
 Identidade pessoal 251-252  
 Quase-memória de Parfit 267-271  
*Ver também:* continuidade psicológica  
*Mental Reality* (Strawson)  
 Sobre a dor 32-33  
 "Mentalés" 88-91, 216-217  
 Metafísica 290-291  
 Identidade pessoal 132-133, 249  
*Metaphysics of Mind, The* (Kenny) 153-154  
 Mill, John Stuart (1806-73)  
 Outras mentes 223-224  
 Mind  
 Acesso privilegiado 25-31  
 Como software 62-64  
 Divisibilidade 63-65  
 Natureza da 13-14  
*Ver também:* problema mente/corpo; outras mentes

*Mind, Brain and the Quantum* (Lockwood) 37n  
*Minds, Brains and Science* (Searle) 87-88, 152-153, 167-168  
 Minsky, Marvin 90-91  
 Monismo  
 Anômalo 14-15, 181-182  
 Definido 291  
 Não-reduutivo 14-15, 291  
*Ver também:* monismo não-reduutivo  
 Monismo anômalo 181-182  
 As objeções de Kim ao 208-212  
 As objeções de Searle ao 194-197  
 Discussão aprofundada 196-199  
 Tese básica 192-194  
 Monismo não-reduutivo 291  
 Dualismo de propriedade 171-172  
*Monty Python* 204-205  
 Moralidade  
 Doutrina da superveniência 163-165  
 "Morning at the Window" (Eliot) 40-41  
 Morte  
 Teoria da resurreição 258-259  
 Visão de Parfit sobre 272-274  
*Mystery of Consciousness, The* (Searle) 170-171, 216-217

**N**

Nagel, Thomas (1937- ) 172  
 Charutos e processos cerebrais 86-88  
 Materialismo cartesiano 278  
 "Physicalism" 83-84  
 "What it is like to be a bat" 158n  
*Naming and Necessity* (Kripke) 96-97  
 Naturalismo 291  
 Biológico 164-172  
 Naturalismo biológico 291  
 Argumento contra micro e macroestrutura 166-169  
 Argumento favorável de Searle 164-167  
*Nature of True Minds, The* (Heil) 30-31  
 Necessidade 287  
 Bicondicional 286  
*De dicto/de re* 291  
 Nemirov, Laurence 150-151  
 Nussbaum, Martha  
 Sobre Aristóteles 207-208

**O**

O'Hearn, Anthony  
 Sobre o eu de Parfit 272-273  
 Ocasionalismo 52-53, 291  
*On Certainty* (Wittgenstein) 241

Ontologia 292  
 De estados mentais 13-14  
 Definição 13-14  
 Reducionista 77-78  
 Relações causais entre eventos 188-189  
 Searle sobre o dualismo 169-172  
 Status de estados mentais 220-221, 246n  
 Outras mentes  
 A abordagem criterial 236-240  
 Aquisição de significado através de termos mentais 235-237  
 Argumento induutivo a partir da analogia 223-225  
 Atribuição eu-e-outro de Strawson 227-229  
 Linguagem privada 220-223  
 Privacidade de estados mentais 220-223  
 Termos mentais substituem comportamento 239-241  
 Wittgenstein ataca o argumento da analogia 241-244  
*Ver também:* privacidade do mental

## P

Paralelismo 52-53  
 Paralelismo psicofísico 292  
 Parfit, Derek (1942- )  
 Cérebros cíndidos 258, 260-261, 280-281  
 Continuidade psicológica de Locke 266-271  
 Dispositivo de transferência de estado cerebral 273-277  
 Peacocke, Christopher (1950- )  
 Perspectiva interna 26-27  
 Pensamento  
 Definição 46-47  
 Dúvida de Descartes 53-58  
 Mentalés 88-91  
 Percebendo outros 67-68  
 Percepções distintas de Descartes 58-64  
 Pepys, Samuel 26-27  
 Percepções 16-17  
 Conteúdo intencional 21-23  
 Representações e apresentações 22-23  
*Perception* (Robinson) 32-33  
 Pessoas não-humanas 206-208  
*Phaedo* (Platão) 41-42  
*Philosophy of Mind: An Introduction* (Graham) 243-244  
 "Physicalism" (Nagel) 83-84  
*Physicalism* (Wilkes) 101-102, 141-142  
 Place, U.T. (1924- )

- Teoria da identidade 72-73  
 Platão (c.428-347 c.C.) 292  
*A República* 41-42  
 Dialogo com Alcefbades 41-45  
 Dualismo 41-45  
 Mito da caverna 162-163  
 Morte de Sócrates 210-212  
*Phaedo* 41-42, 210-211  
 Sobre a alma 43-48, 204-205
- Popper, Karl (1902-94)  
 Materialismo eliminativo 104n  
 Teoria da conexão cerebral 214-215
- Positivismo lógico 292  
 O Círculo de Viena 109-110  
 Verificacão 295
- Pratchett, Terry 23-25
- Privacidade do mental  
 Definições ostensivas 229-235  
 Dor 32-33  
 Experiência 25-28  
 Linguagem pessoal e interna 227-229  
 Problema de outras mentes 221-223  
 Qualia e irredutibilidade 85-88
- Problema mente/corpo  
 E Aristóteles 201-202, 214-215  
 Experiências subjetivas 12-14  
 Interação 106-107  
 Naturalismo biológico 164-166  
 Ocasionismo 52-53  
 Paralelismo 52-53  
 Relação dos fenômenos físicos e mentais 13-15  
 Ver também: dualismo; materialismo; teoria da identidade mente/cérebro
- Propriedades  
 Distintas da substância 40-42  
 Doutrina da superveniência 160-161  
 Mental 171-173
- Proust, Marcel  
*A la recherche du temps perdu* 29-30
- Psicologia  
 Behaviorismo 106-108  
 Bi ciências "duras" 109-110
- Psiche 203-208, 214-218
- Putnam, Hilary (1926- ) 295  
 Capacidades 209-210  
 Causas e efeitos 222-223  
 Comportamento e sintomas 119-120  
 Sobre Aristóteles 207-208  
 Superespartanos 123-125
- Q**  
 Qualia 292-293  
 Argumento das cabeças de homúnculos 144-146  
 Behaviorismo 120
- Estados mentais 19-21  
 Experiências privadas 27-28  
 Invertida 146-147  
 Teoria da identidade 85-87  
 "What Mary Didn't Know" 147-151
- "Qualia epifenomenica" (Jackson) 147-148
- Quine, W.V.O. 217-218
- R**
- Racionalidade  
 Teoria da identidade 93-94
- Ramsey, Frank Plumpton (1903-30)  
 "Ramsey sentence" 116-117, 142-143, 294
- Rediscovery of the Mind, The* (Searle) 95-96, 152-153
- Reducionismo  
 Abordagem criterial 238-239  
 Analítico 293  
 Behaviorismo analítico 108, 221-223  
 Estar consciente de 173  
 Funcionalismo 142-137  
 Identidade pessoal 250-253, 278-283  
 Monismo anômalo de Davidson 193-194  
 Ontológico 77, 293  
 Searle rejeita o materialismo 169-172, 198-199  
 Teoria da identidade 74-75, 179n  
 Teoria da identidade espécime-espécime 93-97  
 Teoria da identidade tipo-tipo 77-79
- Reid, Thomas (1710-96)  
 Objeta a identidade pessoal de Locke 265-266
- República, A* (Platão) 41-42  
 Mito da caverna 162-163
- Robinson, Howard  
*Perception* 32-33
- Robôs. Ver: inteligência artificial
- Russell, Bertrand (1872-1970)  
 Conteúdo intencional 23-25  
 Conhecimento por descrição 150-151  
 E Wittgenstein 89-90
- Ryle, Gilbert (1900-76)  
 Behaviorismo "brando" 116-119  
 Behaviorismo 126-127n  
 Corpo e não-corpo no *The Concept of Mind* 45-46  
 Disposições 119-120  
 Negação de estados mentais subjetivos 120-121  
*The Concept of Mind* 116-117
- S**
- Sacks, Oliver  
*An Anthropologist on Mars* 37n, 157n  
*The Man Who Mistook His Wife for a Hat* 126n
- Sartre, Jean Paul 134
- Searle, John (1932- ) 25-26  
 Argumento contra micro e macroestruturas 166-169  
 Argumento do quarto chinês 151-154  
 Computadores e intencionalidade 25-26  
 Direção de ajuste 20-21  
 Dualismo de propriedade 169-172  
 Homúnculos e funcionalismo 154-156  
 Intencionality 88-89, 167-168  
 Localização da sede e visão no cérebro 167-169  
*Minds, Brains and Science* 87-88, 152-153, 167-168, 215-216  
 Modelos de mentes tecnológicas 136  
 Naturalismo biológico 164-167, 291  
 Objecções ao monismo anômalo 194-197  
 Rejeição do materialismo redutivo 169-172  
*Searle and his Critics* 169-170  
*The Mystery of Consciousness* 170-171, 216-217  
*The Rediscovery of the Mind* 95-96, 152-153  
*Searle and his Critics* 169-170
- Sede  
 Localização no cérebro 167-169
- Sensações 16-17  
 Associação com emoções 19-20  
 Atomística 33-34  
 Auto-atribuições 32-33  
 Conteúdo intencional 20-23  
 Dúvida de Descartes 53-54  
 Entre ser físico e alma 50-53  
 Fenomenologia 19-20  
 Localizações corporais 16-19  
 Sito no cérebro 168-170  
 Ter consciência de 18-19
- Sentidos  
 Ver também: cores
- Ser pessoa versus ser humano 132-136
- Sonhos  
 Malcolm sobre 126n
- Sherrington, Sir Charles (1857-1952)  
 Mente como uma troca telefônica 136
- Shipemaker, Sydney (1931- )  
 Circularidade 116-117  
 Sobre o funcionalismo 142-143  
 Teoria da continuidade física 257-258
- Significado natural 90-91

- Silogismos 43  
 Sintaxe 152-153  
 Skinner, B.F.  
     Behaviorismo 107-108  
 Smart, J.J.C. (1920- )  
     Teoria da identidade 72-73  
 "Sobredade". Ver: intencionalidade  
 Sócrates ver: Platão  
 Solipsismo 183-184, 294  
 Spinoza, Baruch (1622-1677)  
     Causa de ter consciência de 212-213  
     Objeções aos testes cartesianos da alma 49-50  
 Strawson, Galen (1952- ) 88-89, 243-244  
     Mental Reality 32-33  
     Sobre Wittgenstein e behaviorismo 238-239  
 Strawson, Peter F. (1919-2006)  
     *Individuals* 202, 241-244  
     Intencionalidade 88-89  
     Linguagem 133  
     Mente/corpo 202  
     Objeções ao argumento 243-244  
 Substância  
     Aristóteles 202-203  
     Definida 294  
     Distinta de propriedades 39-41  
     E a alma 39-41  
 Suficiência  
     Bicondicional 286  
 Swinburne, Richard (1934- )  
     A alma 279-282  
 System of Nature, The (d' Holbach) 73-74
- T**  
 Tautologia 294  
 Telepatia 246n  
 Teoria da conexão cerebral 214-215  
 Teoria da continuidade física 252-255  
     Argumento de Unger a favor da continuidade psicológica 277-279  
     Cérebros cíndidos 258-262, 273-277  
     Homem e pessoa de Locke 254-257  
     Identidade sem 273-277  
     Leibniz sobre 263n  
     Transplantes de cérebros 254-258  
 Teoria da identidade espécime-espécime  
     Argumento reducionista 93-97  
     Comparada à tipo-tipo 78-81  
     Doutrina da superveniência 160-162  
     Objeção de Kripke 96-97, 99-101  
 Teoria da identidade mente/cérebro.  
     Ver: teoria da identidade  
 Teoria da identidade tipo-tipo 76-77
- Comparada à espécime-espécime 78-81  
 E funcionalismo 132-133, 136  
 Objeções de Kripke 96-97, 99-101  
 Reducionismo 77-79  
 Teoria pictórica do significado 89-90  
 Teoria representacionista 217-218  
 Ter consciência de  
     De sensações 18-19  
     Estados não sensoriais 18-20  
 Tese da translucidez 27-31  
 Teste de Turing 120-122  
 Thales (c. 600 a.C.)  
     Água como elemento básico 76-77  
 Transcendentalismo 209-210, 212-213  
 Turing, Alan (1912-54)  
     Máquina de Turing 137-141
- U**  
 Ulysses (Joyce) 48-49  
 Unger, Peter  
     Continuidade psicológica e cérebro 276-279
- V**  
 Valor de verdade  
     Expressões de dor 239-240  
 Verificacionismo 295  
     Behaviorismo "brando" de Ryle 116-118  
 Vesey, Godfrey 210-212  
 Visão  
     Localização no cérebro 167-169
- W**  
 Watson, J.B. 107-108  
 Watson, J.D. 76-77  
 Weisskrantz, Larry 37n  
 Wernicke, Carl 78-79, 208-209  
 "What is it like to be a bat?" (Nagel) 158n  
 "What Mary Didn't Know" (Jackson) 147-151  
 Wiggins, David (1933- )  
     Cérebros cíndidos 258  
 Wilkes, Katherine V. (1946-2003)  
     Doutrina aristotélica do intelecto ativo 213-214  
     Hormônios e funcionalismo 155-156  
     Physicalism 101-102, 141-142  
     Ter consciência de 211-212  
 Williams, Bernard (1929-2003) 281-282
- Identidade pessoal continuada 251-252  
 Wilson, M.D. 59-60  
 Winnie the Pooh 203  
 Wittgenstein, Ludwig (1889-1951)  
     Abordagem criterial 236-240  
     Aquisição de significado através de termos mentais 235-237  
     Ataca argumento da analogia 224-227  
     Ataca linguagem logicamente privada 227-229  
     Atribuições de primeira pessoa 31-33  
     Behaviorismo analítico 183-185, 238-239  
     Besouro-na-caixa 27-28, 146, 225-226  
     Definições ostensivas 228-236  
     Divisibilidade da mente 64-65  
     Palácio mereológico 216-217  
     Interpretação e sintaxe de imagens 232-235  
     On Certainty 241  
     Sobre ser humano 202  
     Teoria pictórica do significado 89-90  
     The Blue Book 229-230, 233-234  
     The Brown Book 229-230  
     Tractatus Logico-Philosophicus 89-90  
     Uso público de símbolos 154-155  
     Volume de experiência 20-21
- Z**  
 Zumbis 123-124, 280-281

|       |            |
|-------|------------|
| lote  | 325088     |
| data  | 23.11.2014 |
| libr. | Global     |
| preço | R\$ 59,22  |