

# O PROBLEMA MENTE-CORPO



**Jerry Alan Fodor**  
(1935-2017)  
(Rutgers University)

Original: FODOR, J.A. (1981), “The mind-body problem”, *Scientific American* 244(1): 124-32, 148 (janeiro).

Disponível online em:

<http://philosophyfaculty.ucsd.edu/faculty/rarneson/Courses/fodorphil1.pdf>

O presente texto é uma reimpressão, com pequenas modificações, da tradução feita por Saulo de Freitas Araujo (Universidade Federal de Juiz de Fora), disponível no sítio Filosofia da Mente no Brasil: <http://www.filosofiadamente.org/images/stories/textos/fodor.doc> .

As páginas indicadas no centro do rodapé e as figuras correspondem ao original em inglês (cada página do original cobre duas ou três páginas da tradução, que são distinguidas por “a”, “b” e “c”). Uma numeração corrida da presente versão é apresentada à direita no rodapé, entre colchetes: [1]...

Reimpressão preparada por Osvaldo Pessoa Jr. para a disciplina  
TCFC3 – Filosofia das Ciências Neurais, FFLCH-USP, São Paulo, 2014.

# O Problema Mente-Corpo

*Seria possível as máquinas de calcular sentirem dor, os marcianos terem expectativas e os espíritos desencarnados pensarem? A moderna abordagem funcionalista na psicologia levanta a possibilidade lógica de isso acontecer.*

por Jerry A. Fodor

A moderna filosofia da ciência tem se dedicado, em grande parte, à descrição formal e sistemática das práticas bem sucedidas dos cientistas. O filósofo não tenta ditar como a investigação e os argumentos científicos devem ser conduzidos. Ao invés disso, ele procura enumerar os princípios e práticas que têm contribuído para a boa ciência. A análise tem ajudado a tornar mais evidente a natureza da confirmação, a estrutura lógica das teorias científicas, as propriedades formais dos enunciados que expressam leis e a questão da existência real das entidades teóricas.

Só muito recentemente é que os filósofos começaram a se interessar seriamente pelos princípios metodológicos da psicologia. As explicações psicológicas do comportamento referem-se, de maneira ampla, à mente e aos seus estados, operações e processos. A dificuldade filosófica surge quando procuramos determinar, em uma linguagem não ambígua, o que essas referências implicam.

As filosofias tradicionais da mente podem ser divididas em duas grandes categorias: as teorias dualistas e as teorias materialistas. Segundo a abordagem dualista, a mente é uma substância não-física. Para as teorias materialistas, o mental não é diferente do físico; na verdade, todos os estados, propriedades, operações e processos mentais são, em princípio, idênticos a estados, propriedades, operações e processos físicos. Alguns materialistas, conhecidos como behavioristas

[ou comportamentalistas], afirmam que toda discussão sobre causas mentais pode ser eliminada da linguagem da psicologia e substituída pela discussão dos estímulos ambientais e das respostas comportamentais. Outros materialistas, os teóricos da identidade, defendem a existência de causas mentais e a identidade dessas com eventos neurofisiológicos no cérebro.

Nos últimos 15 anos, surgiu uma filosofia da mente chamada de funcionalismo. Essa filosofia, que não é nem dualista nem materialista, é o resultado de uma reflexão filosófica sobre os desenvolvimentos da inteligência artificial, da teoria computacional, da linguística, da cibernética e da psicologia. Todos esses campos, conhecidos coletivamente como as ciências cognitivas, possuem em comum um certo nível de abstração e uma preocupação com sistemas e processos de informação. O funcionalismo, que é uma tentativa de fornecer uma explicação filosófica desse nível de abstração, reconhece a possibilidade de sistemas tão diversos como os seres humanos, as máquinas de calcular e os espíritos desencarnados poderem ter estados mentais. Segundo a visão funcionalista, a psicologia de um sistema não depende da matéria a partir da qual ela é feita (células vivas, energia mental ou espiritual), mas sim do modo como ela é arranjada. 'Funcionalismo' é um conceito difícil e uma das maneiras de entendê-lo é rever as deficiências das filosofias dualistas e materialistas que ele pretende substituir.

O principal inconveniente do dualismo é o seu fracasso em explicar adequadamente a causação mental. Se a mente é não-física, ela não ocupa posição no espaço físico. Como, então, pode uma causa mental provocar um efeito comportamental que tem uma posição no espaço? Em outras palavras, como pode o não-físico dar origem ao físico, sem violar as leis da conservação de massa, de energia e do momentum?

O dualista poderia responder que o problema de como uma substância imaterial pode causar eventos físicos não é mais obscuro do que o problema de como um evento físico pode causar outro. Entretanto, há uma diferença importante: existem muitos casos evidentes de causação física, mas nenhum caso evidente de causação não-física. A interação física é algo com que os filósofos, como todas as outras pessoas, têm que conviver. A interação não-física, por outro lado, pode ser apenas um artefato [efeito espúrio] da concepção imaterialista do mental. Hoje em dia, a maioria dos filósofos concorda que nenhum argumento demonstrou com sucesso por que a causação mente-corpo não deve ser considerada como uma espécie de causação física.

O dualismo é incompatível também com as práticas dos psicólogos. O psicólogo aplica frequentemente os métodos experimentais das ciências físicas ao estudo da mente. Se os processos mentais fossem de um tipo diferente dos processos físicos, não haveria razão para esperar que esses métodos funcionassem no domínio do mental. Para justificar seus métodos experimentais, muitos psicólogos procuraram uma alternativa ao dualismo.

Na década de 20, John B. Watson, da John Hopkins University, fez a sugestão radical de que o comportamento não tem causas mentais. Ele considerou o comportamento de

um organismo como sendo suas respostas observáveis a estímulos, que seriam as verdadeiras causas do comportamento. Nos 30 anos posteriores, psicólogos como B.F. Skinner, da Harvard University, desenvolveram as ideias de Watson, construindo uma elaborada visão de mundo, na qual o papel da psicologia seria catalogar as leis que determinam as relações causais entre estímulos e respostas. Segundo essa visão do “behaviorismo radical”, o problema de explicar a natureza da interação mente-corpo desaparece, uma vez que tal interação não existe.

O behaviorismo radical sempre soou um pouco paradoxal. A ideia de causação mental está profundamente arraigada em nossa linguagem cotidiana e na maneira pela qual nós compreendemos nossos companheiros e a nós mesmos. As pessoas, por exemplo, atribuem normalmente o comportamento a crenças, ao conhecimento e a expectativas. Brown coloca gasolina no tanque de seu carro porque ele acredita que o carro não andar sem ela. Jones escreve “*achieve*” ao invés de “*acheive*” porque ele conhece a regra de colocar o “i” antes do “e”. Mesmo quando uma resposta comportamental está intimamente relacionada a um estímulo ambiental, os processos mentais frequentemente intervêm. Smith carrega um guarda chuva porque o céu está nublado. O tempo, porém, é apenas uma parte da história. Aparentemente, existem também conexões mentais na cadeia causal: observação e expectativa. As nuvens afetam o comportamento de Smith apenas porque ele as observa e porque elas o induzem a uma expectativa de chuva.

Os apelos a tais casos não afetam o behaviorista radical. Ele dispensa referências a causas mentais, por mais plausível que possam parecer, considerando-as um resíduo de crenças ultrapassadas. Os behavioristas

radicais fazem a previsão de que à medida que os psicólogos forem alcançando um maior entendimento sobre as relações entre estímulos e respostas, eles reconhecerão a enorme possibilidade de se explicar o comportamento sem que sejam postuladas causas mentais.

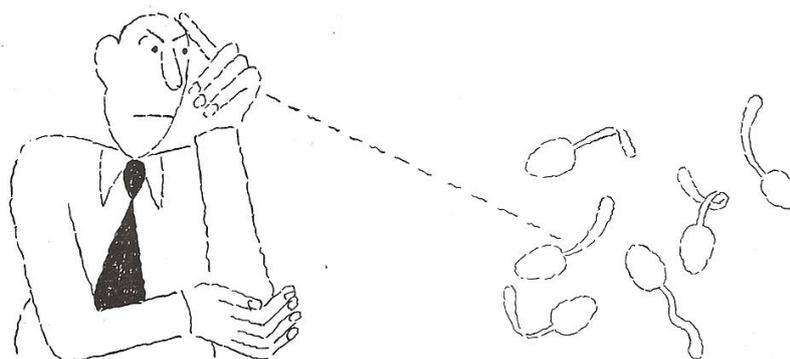
O mais forte argumento contra o behaviorismo é que a psicologia não seguiu esse caminho, mas exatamente o oposto. À medida que a psicologia vem amadurecendo, a estrutura de estados e processos mentais aparentemente necessária para explicar as observações experimentais vem se tornando cada vez mais elaborada. Particularmente no caso do comportamento humano, as teorias psicológicas que satisfazem os princípios metodológicos do behaviorismo radical têm se mostrado em grande parte estéreis, como seria de se esperar, caso os processos mentais postulados sejam reais e causalmente efetivos.

Muitos filósofos foram inicialmente atraídos para o behaviorismo radical porque, apesar dos paradoxos e de todas as dificuldades, ele parecia melhor do que o dualismo. Uma vez que uma psicologia comprometida com substâncias imateriais era inaceitável, os filósofos viraram-se para o

behaviorismo radical porque ele parecia ser a única alternativa em termos de uma filosofia materialista da mente. A escolha, como eles a entendiam, era entre o behaviorismo radical e fantasmas.

No início dos anos 60, os filósofos começaram a duvidar que o dualismo e o behaviorismo radical seriam as únicas abordagens possíveis na filosofia da mente. Já que ambas as teorias não pareciam atraentes, a estratégia correta poderia ser o desenvolvimento de uma filosofia materialista da mente que aceitasse, entretanto, causas mentais. Assim, surgiram duas novas filosofias da mente, a saber, o behaviorismo lógico e a teoria da identidade.

O behaviorismo lógico é uma teoria semântica sobre o significado dos termos mentalistas. Sua ideia básica é que atribuir um estado mental (digamos, sede) a um organismo é o mesmo que dizer que o organismo está disposto a se comportar de um determinado modo (por exemplo, beber, se houver água disponível). De acordo com esse ponto de vista, toda atribuição mental é semanticamente equivalente a um enunciado do tipo “se-então” (denominada hipotético



**O DUALISMO é a filosofia da mente que considera a mente como uma substância não-física. Ele divide tudo o que há no mundo em duas categorias distintas: o mental e o físico. A principal dificuldade do dualismo é o seu fracasso em dar conta adequadamente da interação causal entre o mental e o físico. Não é evidente como uma mente não-física poderia gerar qualquer efeito físico sem violar as leis de conservação de massa, energia e momentum.**

comportamental), que expressa uma disposição comportamental. Por exemplo, o enunciado “Smith está com sede” poderia ser considerado equivalente ao enunciado disposicional “Se houvesse água disponível, então Smith beberia um pouco”. Por definição, um hipotético comportamental não inclui termos mentalistas. A oração iniciada pelo “se” do hipotético refere-se apenas a estímulos e a oração iniciada pelo “então” faz referência somente a respostas comportamentais. Uma vez que os estímulos e as respostas são eventos físicos, o behaviorismo lógico é um tipo de materialismo.

A força do behaviorismo lógico reside no fato de que a tradução da linguagem mentalista para a linguagem de estímulos e respostas fornece uma interpretação daquelas explicações psicológicas em que os efeitos comportamentais são atribuídos a causas mentais. A causação mental é simplesmente a

manifestação de uma disposição comportamental. Mais precisamente, a causação mental é o que acontece quando um organismo tem uma disposição comportamental e a oração iniciada pelo “se” do hipotético comportamental, que expressa uma disposição, é verdadeira. Por exemplo, o enunciado causal “Smith bebeu um pouco d’água porque ele estava com sede” poderia ser considerado semanticamente equivalente ao enunciado “Se houvesse água disponível, então Smith beberia um pouco, e havia água disponível”.

De certa forma, eu simplifiquei muito o behaviorismo lógico, ao assumir que toda atribuição mental pode ser traduzida por um único hipotético comportamental. Na verdade, o behaviorista lógico afirma frequentemente que seria necessário um conjunto aberto (talvez um conjunto infinito) de hipotéticos comportamentais para indicar a disposição comportamental expressa por um termo



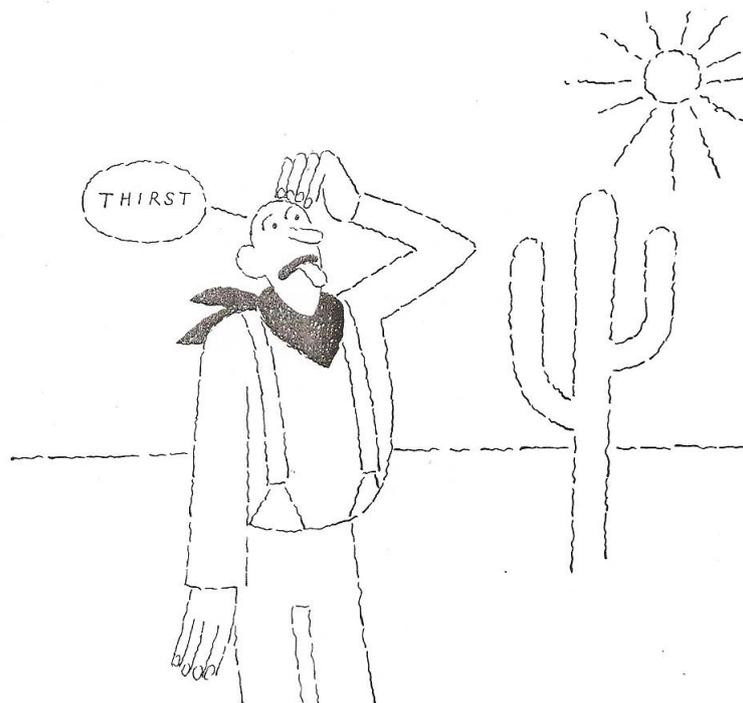
**O BEHAVIORISMO RADICAL é a filosofia da mente que nega a existência da mente e de estados, propriedades, processos e operações mentais. O behaviorista radical acredita que o comportamento não tem causas mentais. Ele considera que o comportamento de um organismo são suas respostas a estímulos. O papel da psicologia é catalogar as relações entre estímulos e respostas.**

mentalista. A atribuição mental “Smith está com sede” poderia também ser satisfeita pelo hipotético “Se houvesse suco de laranja disponível, então Smith beberia um pouco” e também por vários outros hipotéticos. Em todo caso, o behaviorista lógico normalmente não afirma que ele pode realmente enumerar todos os hipotéticos que correspondem a uma disposição comportamental expressa por um termo mentalista. Ele apenas insiste que qualquer termo mentalista pode ser traduzido por hipotéticos comportamentais.

O modo pelo qual o behaviorista lógico tem interpretado um termo mentalista – como, por exemplo, sede – baseia-se na maneira pela qual muitos filósofos têm interpretado uma disposição física, como, por exemplo, fragilidade. A disposição física “O copo é

frágil” é frequentemente considerada equivalente a algo do tipo “Se o copo sofresse alguma batida, então ele quebraria”. Do mesmo modo, a análise da causação mental feita pelos behavioristas lógicos é similar à análise tradicional de um tipo de causação física. O enunciado causal “O copo quebrou porque era frágil” é considerado equivalente a algo do tipo “Se o copo sofresse alguma batida, então ele quebraria, e o copo sofreu uma batida”.

Ao identificar termos mentalistas com disposições comportamentais, o behaviorista lógico colocou os termos mentalistas no mesmo patamar das disposições não-comportamentais das ciências físicas. Trata-se de uma mudança promissora porque a análise das disposições não-comportamentais repousa sobre uma base filosófica relativamente



**O BEHAVIORISMO LÓGICO** é uma tese semântica sobre o que termos mentais significam. O behaviorista lógico defende que termos mentais exprimem disposições comportamentais. Considere o estado mental de estar com sede (*thirst*). O behaviorista lógico defende que a sentença “Smith está com sede” pode ser tomado como sendo equivalente em significado a um enunciado disposicional “Se houvesse água disponível, então Smith a beberia”. A força do behaviorismo lógico é que ele fornece um relato da causação mental: a realização de uma disposição comportamental. Por exemplo, o enunciado causal “Smith bebeu um pouco de água porque estava com sede” pode ser tomado como significando “Se houvesse água disponível, então Smith a beberia, e havia água disponível”.

sólida. Uma explicação que atribui a quebra de um copo à sua fragilidade é certamente algo que mesmo o materialista mais fiel pode aceitar. Ao argumentar que os termos mentalistas são sinônimos dos termos disposicionais, o behaviorista lógico fornece algo que o behaviorista radical não conseguiu, a saber, uma explicação materialista da causação mental.

Em que pese tudo isso, a analogia feita pelo behaviorista lógico entre a causação mental e a causação física só vai até aqui. O behaviorista lógico trata a manifestação de uma disposição como a única forma de causação mental, enquanto as ciências físicas reconhecem outras formas de causação. Há o tipo de causação em que um evento físico causa outro, como no caso da quebra do copo ser atribuída a uma batida por ele sofrida. De fato, as explicações que envolvem a causação entre eventos são presumivelmente mais básicas do que as explicações disposicionais porque a manifestação de uma disposição (a quebra de um copo frágil) envolve sempre a causação entre eventos, e não vice-versa. No domínio do mental, muitos exemplos de causação entre eventos referem-se a um estado mental causando outro. Para esse tipo de causação, porém, o behaviorista lógico não fornece nenhuma análise. Como resultado, ele se vê obrigado a assumir a suposição tácita e implausível de que a psicologia requer uma noção de causação menos robusta do que requerem as ciências físicas.

A causação entre eventos parece ser realmente bastante comum no domínio do mental. As causas mentais ocasionam efeitos comportamentais em virtude de sua interação com outras causas mentais. Por exemplo, ter uma dor de cabeça causa uma disposição de tomar aspirina apenas se houver também o desejo de ficar livre da dor de cabeça, a crença de que a aspirina existe, a crença de que tomar aspirina reduz a dor de cabeça, e assim por diante. Uma vez que os estados mentais interagem para gerar o comportamento, será necessário encontrar uma interpretação das explicações

psicológicas que postule processos mentais: sequências causais de eventos mentais. E é exatamente essa interpretação que o behaviorismo lógico não consegue fornecer.

Tais considerações revelam uma semelhança fundamental entre o behaviorismo lógico e o behaviorismo radical. É bem verdade que o behaviorista lógico, diferentemente do behaviorista radical, reconhece a existência de estados mentais. Entretanto, uma vez que o princípio subjacente ao behaviorismo lógico é que as referências a estados mentais, presentes nas explicações psicológicas, podem ser traduzidas por hipotéticos comportamentais, toda a discussão sobre estados e processos mentais torna-se, num certo sentido, heurística. Os únicos fatos com os quais o behaviorista está realmente comprometido são aqueles que dizem respeito a relações entre estímulos e respostas. Nesse sentido, o behaviorismo lógico é apenas uma versão semântica do behaviorismo radical. Embora o primeiro ofereça uma interpretação da causação mental, a interpretação é pickwickiana [não usual]. Aquilo que de fato não existe, não pode causar nada. E o behaviorista lógico, assim como o behaviorista radical, acredita firmemente que não existem causas mentais.

Uma outra teoria materialista da mente, que se apresenta como alternativa ao behaviorismo lógico, é a teoria da identidade. De acordo com essa teoria, os eventos, estados e processos mentais são idênticos a eventos neurofisiológicos no cérebro, e a propriedade de estar em um certo estado mental (tal como ter uma dor de cabeça ou acreditar que irá chover) é idêntica à propriedade de estar em um certo estado neurofisiológico. A partir disso, é fácil compreender a ideia de que um efeito comportamental poderia, algumas vezes, ter uma cadeia de causas mentais. Isso acontecerá sempre que um efeito comportamental depender da sequência apropriada de eventos neurofisiológicos.

A teoria da identidade do estado central reconhece a possibilidade de as causas mentais interagirem causalmente sem gerarem qualquer efeito comportamental, como, por exemplo, quando uma pessoa pensa por um momento sobre o que ela deveria fazer e então decide não fazer nada. Se os processos mentais são neurofisiológicos, eles devem ter as propriedades causais dos processos neurofisiológicos. Já que esses últimos são presumivelmente processos físicos, a teoria da identidade assegura que o conceito de causação mental é tão rico quanto o conceito de causação física.

A teoria da identidade fornece uma explicação satisfatória do problema da referência dos termos mentalistas presentes nas explicações psicológicas e, assim, tem a aceitação dos psicólogos que estão insatisfeitos com o behaviorismo. O behaviorista sustenta que os termos mentalistas não se referem a nada ou que eles referem-se aos parâmetros das relações estímulo-resposta. Em ambos os casos, a existência das entidades mentais é simplesmente ilusória. O teórico da identidade, por outro lado, argumenta que os termos mentalistas referem-se a estados neurofisiológicos. Sendo assim,



**A TEORIA DA IDENTIDADE DE ESTADO CENTRAL é a filosofia da mente que iguala eventos, estados e processos mentais com eventos neurofisiológicos. A propriedade de estar em um dado estado mental seria idêntica à propriedade de estar em um dado estado neurofisiológico.**

ele pode assumir seriamente o projeto de explicar o comportamento fazendo apelo às suas causas mentais.

A principal vantagem da teoria da identidade é que ela interpreta os construtos explicativos da psicologia ao pé da letra, algo que certamente toda filosofia da mente deveria, se possível, fazer. A teoria da identidade mostra como as explicações mentalistas da psicologia poderiam deixar de ter um sentido meramente heurístico, tornando-se um relato literal da história causal do comportamento. Além disso, uma vez que a teoria da identidade não é uma tese semântica, ela é imune a muitos argumentos que colocam o behaviorismo lógico em dúvida. Um inconveniente do behaviorismo lógico é que a observação “John tem uma dor de cabeça” não parece significar a mesma coisa que o enunciado “John está disposto a se comportar de tal e tal modo”. O teórico da identidade, contudo, pode conviver com o fato de que os enunciados “John tem uma dor de cabeça” e “John está em tal e tal estado cerebral” não são sinônimos. A asserção do teórico da identidade não é que esses enunciados significam a mesma coisa, mas apenas que eles são considerados verdadeiros (ou falsos) pelos mesmos fenômenos neurofisiológicos.

A teoria da identidade pode ser considerada ou como uma doutrina sobre particulares mentais (a dor atual de John ou o medo que Bill tem de animais) ou como uma doutrina sobre universais mentais, ou propriedades (ter uma dor ou ter medo de animais). As duas doutrinas, denominadas, respectivamente, fisicalismo de eventos [ou de ocorrência, instância, espécie; *token physicalism*] e fisicalismo de tipos [*type physicalism*], diferem tanto em termos de

força quanto em termos de plausibilidade. O fisicalismo de eventos sustenta apenas que todos os particulares mentais que calharam de existir são neurofisiológicos, enquanto que o fisicalismo de tipos faz a asserção mais arrebatedora de que todos os particulares mentais possivelmente existentes são neurofisiológicos. O fisicalismo de eventos não descarta a possibilidade lógica de máquinas e espíritos desencarnados terem propriedades mentais. O fisicalismo de tipos rejeita essa possibilidade porque nem as máquinas e nem os espíritos possuem neurônios.

O fisicalismo de tipos não é uma doutrina plausível sobre as propriedades mentais, ainda que o fisicalismo de eventos esteja certo sobre os particulares mentais. O problema com o fisicalismo de tipos é que a constituição psicológica de um sistema parece depender não de seu *hardware*, ou seja, sua composição física, mas sim de seu *software*, isto é, seu programa. Por que o filósofo deveria rejeitar a possibilidade de marcianos constituídos de silício sentirem dor, desde que o silício esteja adequadamente organizado? E por que o filósofo deveria eliminar a possibilidade de máquinas terem crenças, desde que as máquinas estejam corretamente programadas? Se é logicamente possível que marcianos e máquinas possam ter propriedades mentais, então estas últimas não podem ser idênticas a processos neurofisiológicos, independentemente do quanto estes possam ser co-extensivos àquelas.

O que tudo isso revela é que parece haver um nível de abstração, no qual as generalizações da psicologia encaixam-se mais naturalmente. Esse nível de abstração

ignora as diferenças na composição física dos sistemas aos quais as generalizações psicológicas se aplicam. Pelo menos nas ciências cognitivas, o domínio natural da teorização psicológica parece incluir todos os sistemas que processam informação. O problema com o fisicalismo de tipos é que há possíveis sistemas de processamento de informação com a mesma constituição psicológica dos seres humanos, mas com uma organização física diferente. Em princípio, todos os tipos de coisas fisicamente diferentes poderiam ter o software humano.

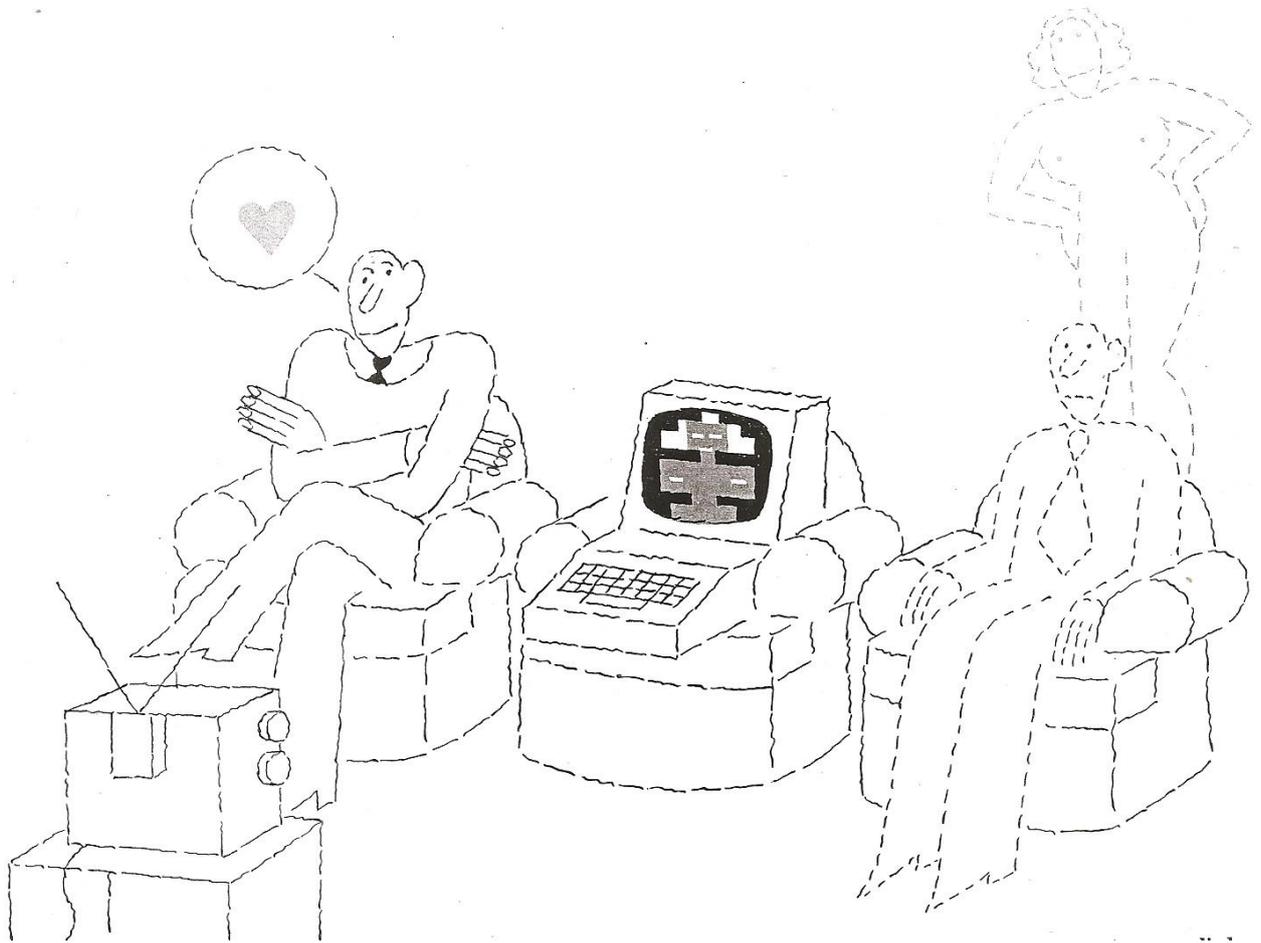
Essa situação exige uma explicação relacional das propriedades mentais que abstraíam-nas da estrutura física de seus portadores. Em que pese as objeções que eu apresentei anteriormente ao behaviorismo lógico, ele estava pelo menos no caminho certo, ao oferecer uma interpretação relacional das propriedades mentais: ter uma dor de cabeça é estar disposto a exibir um certo padrão de relações entre os estímulos que se encontram e as respostas que se exibem. Entretanto, se ter uma dor de cabeça for isso mesmo, não há razão, em princípio, para que apenas as cabeças que são fisicamente semelhantes às nossas possam doer. De fato, de acordo com o behaviorismo lógico, é uma verdade necessária que qualquer sistema que tenha nossas contingências estímulo-resposta também tenha nossas dores de cabeça.

Tudo isso surgiu há 10 ou 15 anos atrás [meados dos anos 1960], como um sério dilema para o programa materialista na filosofia da mente. Por um lado, o teórico da identidade (mas não o behaviorista lógico) estava certo no que diz respeito ao caráter causal das interações da mente e do corpo. Por outro lado, o behaviorista lógico (mas não o teórico da identidade) estava certo no que

diz respeito ao caráter relacional das propriedades mentais. O funcionalismo aparentemente conseguiu resolver o dilema. Ao enfatizar a distinção que a ciência da computação traça entre o *hardware* e o *software*, o funcionalista pode dar sentido tanto ao caráter causal quanto ao caráter relacional do mental.

A intuição por trás do funcionalismo é que o que determina o tipo psicológico ao qual um particular mental pertence é o papel causal do particular na vida mental do organismo. Indivuação funcional é diferenciação com respeito ao papel causal. Uma dor de cabeça, por exemplo, é identificada com o tipo de estado mental que, entre outras coisas, causa uma disposição para tomar aspirina em pessoas que acreditam que a aspirina alivia uma dor de cabeça, causa um desejo de se livrar da dor que se está sentindo, frequentemente causa alguém que fala português a dizer coisas como “Estou com dor de cabeça”, e surge a partir de excesso de trabalho, fadiga ocular e tensão. Esta lista presumivelmente não está completa. Mais será conhecido sobre a natureza de uma dor de cabeça, à medida que pesquisas em psicologia e fisiologia descobrirem mais sobre o seu papel causal.

O funcionalismo elabora o conceito de ‘papel causal’ de um modo tal que um estado mental pode ser definido por suas relações causais com outros estados mentais. Nesse sentido, o funcionalismo é completamente diferente do behaviorismo lógico. Outra grande diferença é que o funcionalismo não é uma tese reducionista. Ele não prevê, nem mesmo em princípio, a eliminação dos conceitos mentalistas do aparato explicativo das teorias psicológicas.



**O FUNCIONALISMO** é a filosofia da mente baseada na distinção que a ciência da computação traça entre o hardware do sistema, ou composição física, e seu software, ou programa. A psicologia de um sistema como um ser humano, uma máquina ou um espírito desencarnado não dependeria do material de que é feito o sistema (neurônios, diodos ou energia espiritual), mas do modo como esse material é organizado. O funcionalismo não descarta a possibilidade, o quão remota que possa ser, de sistemas mecânicos e etéreos terem estados e processos mentais.

A diferença entre o funcionalismo e o behaviorismo lógico é trazida à tona pelo fato de o funcionalismo ser totalmente compatível com o fisicalismo de eventos. O funcionalista não ficaria desconcertado se fosse demonstrado que os eventos cerebrais são as únicas coisas dotadas das propriedades funcionais que definem os estados mentais. De fato, a maioria dos funcionalistas espera que esse será o rumo tomado.

Uma vez que o funcionalismo reconhece que os particulares mentais podem ser físicos, ele é compatível com a ideia de que a causação mental é uma espécie de causação física. Em outras palavras, o funcionalismo tolera a solução materialista para o problema mente-corpo fornecida pela teoria da identidade do estado central. É possível para o funcionalista afirmar tanto que as propriedades mentais são definidas tipicamente em termos de suas relações quanto que as interações da mente e do corpo são tipicamente causais, qualquer que seja a força da noção de causalidade requerida pelas explicações psicológicas. O behaviorista lógico pode aceitar somente a primeira asserção, enquanto que o fisicalista de tipos apenas a segunda. Como consequência, o funcionalismo parece capturar as melhores características das alternativas materialistas ao dualismo. Não chega a ser uma surpresa que o funcionalismo venha se tornando cada vez mais popular.

As máquinas fornecem bons exemplos de duas ideias centrais do funcionalismo: a ideia de que os estados mentais são interdefinidos e a ideia de que eles podem ser realizados por

vários sistemas. A ilustração da página 131 [cuja tabela está redesenhada abaixo] contrasta uma máquina de coca-cola behaviorista com uma mentalista. Ambas as máquinas liberam uma coca-cola por 10 cents de dólar. (O preço não foi afetado pela inflação) Os estados das máquinas são definidos por referência aos seus papéis causais, mas apenas a máquina da esquerda satisfaria o behaviorista. Seu estado único (*S0*) é completamente especificado em termos de estímulos e respostas. *S0* é o estado em que a máquina está se, e somente se, dado uma moeda de 10 cents como entrada [*input*], ela liberar uma coca-cola como saída [*output*].

A máquina da direita tem estados interdefinidos (*S1* e *S2*), que são característicos do funcionalismo. *S1* é o estado em que a máquina está se, e somente se, (1) dada uma moeda de 5 cents, ela não liberar nada e passar para *S2*, e (2) dada uma moeda de 10 cents, ela liberar uma coca-cola e ficar em *S1*. *S2* é o estado em que a máquina está se, e somente se, (1) dada uma moeda de 5 cents, ela liberar uma coca-cola e passar para *S1*, e (2) dada uma moeda de 10 cents, ela liberar uma coca-cola e uma moeda de 5 cents e passar para *S1*. *S1* e *S2* equivalem conjuntamente às seguintes operações: se a máquina recebe uma moeda de 10 cents, libera uma coca-cola; se ela receber uma moeda de 10 e uma de 5 cents, ela libera uma coca-cola e uma moeda de 5 cents; se ela recebe apenas uma moeda de 5 cents, ela aguarda uma segunda moeda de 5 cents.

Máquina behaviorista

	Estado <i>S0</i>
Entrada de 10 cents	Libera uma coca-cola

Máquina funcionalista

	Estado <i>S1</i>	Estado <i>S2</i>
Entrada de 5 cents	Não sai nada e vai para <i>S2</i>	Libera uma coca-cola e vai para <i>S1</i>
Entrada de 10 cents	Libera uma coca-cola e fica em <i>S1</i>	Libera uma coca e mais 5 cents e vai para <i>S1</i>

Uma vez que *S1* e *S2* são definidos por enunciados hipotéticos, eles podem ser vistos como disposições. Entretanto, eles não são disposições comportamentais porque as consequências de uma entrada para uma máquina em *S1* ou *S2* não são especificadas somente em termos da saída da máquina. Ao invés disso, as consequências também envolvem os estados internos das máquinas.

De acordo com a minha descrição das máquinas behaviorista e mentalista de coca-cola, não há restrições sobre sua constituição física. Qualquer sistema, cujos estados mantenham as relações adequadas com entradas, saídas e outros estados, poderia ser uma dessas máquinas. Sem dúvida, é razoável esperar que um tal sistema seja constituído de rodas, alavancas e diodos (fiscalismo de eventos para máquinas de coca-cola), assim como também é razoável esperar que nossas mentes sejam comprovadamente neurofisiológicas (fiscalismo de eventos para seres humanos).

Em que pese tudo isso, a descrição do software de uma máquina de coca-cola não exige logicamente rodas, alavancas e diodos, da mesma maneira que a descrição do software da mente não exige logicamente neurônios. Para o funcionalismo, uma máquina de coca-cola com estados *S1* e *S2* pode ser feita de ectoplasma, caso tal substância exista e seus estados tiverem as propriedades causais adequadas. O funcionalismo não exclui a possibilidade de haver máquinas de coca-cola desencarnadas, da mesma forma como também não exclui a possibilidade de haver mentes desencarnadas.

Dizer que *S1* e *S2* são interdefinidos e realizáveis por diferentes tipos de hardware não significa, é claro, dizer que uma máquina de coca-cola possui uma mente. Embora a interdefinição e a especificação funcional sejam características típicas dos estados mentais, elas não são suficientes para garantir a qualidade do mental. O que ainda falta é uma questão que será abaixo discutida.

Alguns filósofos suspeitam do funcionalismo porque ele parece muito simples. Já que o funcionalismo permite a individuação dos estados através de seu papel causal, ele parece permitir uma explicação trivial de qualquer evento observado *E*, ou seja, ele parece postular um “causador de *E*” [*E-causer*]. Por exemplo, o que faz as válvulas de uma máquina abrirem? Ora, a operação de um abridor de válvulas. E o que é um abridor de válvulas? Ora, qualquer coisa que tenha a propriedade, funcionalmente definida, de causar a abertura de válvulas.

Na psicologia, esse tipo de problema aparece frequentemente nas teorias que efetivamente postulam homúnculos, com as mesmas capacidades intelectuais que o teórico queria explicar. Isso acontece, por exemplo, quando, para explicar a percepção visual, simplesmente postulam-se mecanismos psicológicos que processam a informação visual. O behaviorista tem frequentemente acusado o mentalista, algumas vezes com razão, de cultivar esse tipo de pseudo-explicação que faz uma petição de princípio. Se os estados mentais funcionalmente definidos forem desempenhar um papel importante nas teorias psicológicas, essa acusação terá que ser respondida.

A acusação não é de falsidade, mas sim de trivialidade. Não pode haver dúvidas de que é um abridor de válvulas que abre válvulas, e é provável que a percepção visual seja mediada pelo processamento de informação visual. A acusação é de que tais possíveis explicações funcionais são meras obviedades. O funcionalista pode responder a essa objeção permitindo a introdução de construtos teóricos funcionalmente definidos apenas quando existirem mecanismos capazes de desempenhar a função e quando ele tiver alguma noção de como seriam tais mecanismos. Uma maneira de impor essa exigência é identificar os processos mentais que a psicologia postula com as operações da restrita classe de possíveis computadores, chamados de máquinas de Turing.

Uma máquina de Turing pode ser informalmente caracterizada como um mecanismo com um número finito de estados do programa. As entradas e saídas da máquina são escritas em uma fita, que é dividida em quadrados, cada um deles contendo um símbolo de um alfabeto finito. A máquina escaneia a fita um quadrado de cada vez. Ela pode apagar o símbolo de um quadrado e imprimir um outro em seu lugar. Ela pode executar apenas as seguintes operações mecânicas: escanear, apagar, imprimir, mover a fita [para a esquerda ou para a direita] e mudar de estado.

Os estados do programa da máquina de Turing são definidos somente em termos dos símbolos de entrada e saída da fita, as operações elementares e os outros estados do programa. Cada estado do programa é funcionalmente definido, portanto, pelo papel que ele assume na operação geral da máquina. Uma vez que o papel funcional de um estado depende de sua relação com outros estados e também com as entradas e saídas, o caráter relacional do mental fica preservado na versão máquina-de-Turing do funcionalismo. Já que a definição de um estado do programa nunca se refere à estrutura física do sistema que roda o programa, a versão máquina-de-Turing do funcionalismo também preserva a ideia de que o caráter de um estado mental é independente de sua realização física. Um ser humano, uma sala cheia de pessoas, um computador e um espírito desencarnado seriam todos uma máquina de Turing, se eles operassem de acordo com um programa de uma máquina de Turing.

A proposta é restringir a definição funcional dos estados psicológicos àqueles que puderem ser expressos em termos dos estados do programa de uma máquina de Turing. Se esta restrição puder ser implementada, ela garantirá a compatibilidade das teorias psicológicas com as demandas dos mecanismos. Uma vez que as máquinas de Turing são dispositivos muito simples, elas são, em princípio, bastante fáceis de serem construídas. Consequentemente, ao formular

uma explicação psicológica como um programa de uma máquina de Turing, o psicólogo assegura que sua explicação é mecanística, muito embora ele não especifique o hardware responsável pela execução do mecanismo.

Existem muitos tipos de mecanismos computacionais além das máquinas de Turing. Assim, a formulação de uma teoria psicológica funcionalista na notação da máquina de Turing oferece apenas uma condição suficiente para que a teoria seja mecanicamente realizável. O que torna essa condição interessante, entretanto, é que a simples máquina de Turing pode realizar muitas tarefas complexas. Embora as operações elementares da máquina de Turing sejam restritas, as interações das operações capacitam a máquina a realizar qualquer computação bem definida sobre símbolos discretos.

Uma importante tendência nas ciências cognitivas é tratar a mente sobretudo como um dispositivo que manipula símbolos. Se um processo mental pode ser funcionalmente definido como uma operação sobre símbolos, existe uma máquina de Turing capaz de realizar a computação e uma variedade de mecanismos para operar a máquina de Turing. Quando a manipulação de símbolos é importante, a máquina de Turing faz uma conexão entre a explicação funcional e a explicação mecanística.

A redução de uma teoria psicológica a um programa para uma máquina de Turing é um modo de exorcizar os homúnculos. A redução assegura que não foi postulada nenhuma operação além daquelas capazes de serem realizadas por um mecanismo conhecido. É óbvio que o psicólogo normalmente não pode especificar a redução para cada um dos processos funcionalmente definidos em todas as teorias que ele leva a sério. Na prática, o argumento normalmente vai na direção contrária: se a postulação de uma operação mental é essencial para alguma explicação psicológica, o teórico tende a supor que deve haver um programa para uma máquina de Turing que realizará a operação.

As “caixas pretas”, que são comuns nos fluxogramas feitos pelos psicólogos, servem para indicar processos mentais postulados, que estão à espera de uma especificação em termos de uma máquina de Turing. Mesmo assim, a possibilidade em princípio de tais reduções serve como uma restrição metodológica sobre a teorização psicológica, no sentido de determinar quais definições funcionais serão permitidas e como seria a situação de se saber que tudo aquilo que possivelmente poderia exigir explicação já foi explicado.

Tal é a origem, a proveniência e a promessa do funcionalismo contemporâneo. Mas quais são seus resultados? Essa questão não é fácil de responder porque uma boa parte do que está acontecendo atualmente na filosofia da mente e nas ciências cognitivas é uma exploração do alcance e dos limites das explicações funcionalistas do comportamento. Contudo, eu apresentarei uma breve visão geral.

Uma objeção óbvia ao funcionalismo como teoria da mente é que a definição funcionalista não se limita a estados e processos mentais. Catalisadores, máquinas de coca-cola, abridores de válvula, apontadores, ratoeiras e ministros da fazenda são, de um maneira ou de outra, conceitos funcionalmente definidos, mas nenhum deles é um conceito mental tal como dor, crença e desejo. O que é, então, que caracteriza o mental? E seria possível incluí-lo em um modelo funcionalista?

A visão tradicional na filosofia da mente afirma que os estados mentais distinguem-se por possuírem o que é chamado de conteúdo qualitativo ou conteúdo intencional. Primeiramente, eu discutirei o conteúdo qualitativo.

Não é fácil dizer o que é o conteúdo qualitativo. De fato, segundo algumas teorias, não é possível sequer dizer o que ele é porque ele não pode ser conhecido por descrição, mas apenas através da experiência direta. Apesar disso, eu tentarei descrevê-lo. Tente imaginar-se olhando para uma parede em branco através de um filtro vermelho. Agora, troque

o filtro vermelho por um verde e deixe o resto exatamente como estava antes. Alguma coisa relativa ao caráter de sua experiência se altera quando o filtro é trocado, e é esse tipo de coisa que os filósofos chamam de conteúdo qualitativo. Eu não me sinto inteiramente confortável com essa definição do conteúdo qualitativo, mas trata-se de um tema com o qual muitos filósofos não se sentem confortáveis.

A razão pela qual o conteúdo qualitativo representa um problema para o funcionalismo é evidente. O funcionalismo está comprometido com a definição dos estados mentais em termos de suas causas e seus efeitos. Parece possível, no entanto, que dois estados mentais possam ter as mesmas relações causais, diferindo, porém, em seu conteúdo qualitativo. Deixe-me ilustrar isso com o clássico enigma do espectro invertido.

Aparentemente, é possível imaginar dois observadores semelhantes em todos os aspectos psicológicos relevantes, com a única exceção de que as experiências subjetivas que possuem o conteúdo qualitativo do vermelho para um observador teriam o conteúdo qualitativo do verde para o outro observador. O comportamento de ambos não revela a diferença porque todos os dois vêem um tomate maduro e um pôr do sol flamejante como sendo de cor semelhante e chamam essa cor de “vermelho”. Além disso, a conexão causal entre suas experiências (qualitativamente distintas) e seus outros estados mentais também poderiam ser idênticas. Talvez ambos pensem na Chapeuzinho Vermelho quando vêem tomates maduros, sintam-se deprimidos quando vêem a cor verde, e assim por diante. Parece que qualquer coisa que pudesse ser incluída na noção de papel causal de suas experiências poderia ser compartilhada por eles, e, apesar disso, o conteúdo qualitativo das experiências poderia ser totalmente diferente. Se isto for possível, então a abordagem funcionalista não funciona para os estados mentais que possuem conteúdo qualitativo. Se uma pessoa está tendo uma experiência do verde enquanto

outra está tendo uma do vermelho, elas certamente devem estar em estados mentais diferentes.

O exemplo do espectro invertido é mais do que um enigma verbal. A posse de conteúdo qualitativo é considerada como um dos principais fatores para que um estado mental seja consciente. Muitos psicólogos inclinados a aceitar o quadro funcionalista estão preocupados com o fato de o funcionalismo não ter muita coisa a dizer sobre a natureza da consciência. Os funcionalistas têm feito algumas tentativas ingênuas de convencer seus colegas e a si mesmo de esquecerem essa preocupação, mas não têm obtido, segundo meu ponto de vista, muito sucesso. (Por exemplo, talvez seja errado pensar que é possível imaginar como seria um espectro invertido) No presente momento, o problema do conteúdo qualitativo representa uma séria ameaça para a afirmação de que o funcionalismo pode oferecer uma teoria geral do mental.

No que diz respeito ao conteúdo intencional dos estados mentais, o funcionalismo tem se saído muito melhor. De fato, é aqui que se encontram as maiores realizações da ciência cognitiva. Dizer que um estado mental tem conteúdo intencional é dizer que ele tem certas propriedades semânticas. Por exemplo, para Enrico

acreditar que Galileo era italiano parece necessário haver uma relação ternária entre Enrico, uma crença e uma proposição, que é o conteúdo da crença (a saber, a proposição que Galileo era italiano). Em particular, é uma propriedade essencial da crença de Enrico que ela refere-se a Galileo (e não a Newton, por exemplo) e que ela é verdadeira se, e somente se, Galileo era realmente italiano. Os filósofos não concordam sobre a maneira como essas considerações se encaixam, mas há um acordo geral de que as crenças envolvem propriedades semânticas, tais como expressar uma proposição, ser verdadeiro ou falso e referir-se a uma coisa ao invés de outra.

É importante compreender as propriedades semânticas das crenças porque as teorias nas ciências cognitivas referem-se em grande parte às crenças que os organismos têm. Teorias da aprendizagem e da percepção, por exemplo, são explicações de como o conjunto de crenças que um organismo tem é determinado pelo caráter de suas experiências e de sua herança genética. A abordagem funcionalista dos estados mentais não oferece, por si própria, os *insights* necessários. Ratoeiras são funcionalmente definidas, mas ratoeiras não expressam proposições e não são verdadeiras ou falsas.

Existe pelo menos uma outra coisa, além de um estado mental, que tem conteúdo intencional: um símbolo. Assim como os pensamentos, os símbolos parecem referir-se

a coisas. Se alguém diz “Galileo era italiano”, sua afirmação, como a crença de Enrico, expressa uma proposição sobre Galileo, que é verdadeira ou falsa, dependendo da pátria de Galileo. Esse paralelo entre o simbólico e o mental subjaz à tradicional busca por um tratamento unificado da linguagem e da mente. A ciência cognitiva está agora tentando oferecer tal tratamento.

O conceito básico é simples, porém notável. Assuma que existem símbolos mentais (representações mentais) e que os símbolos mentais tenham propriedades semânticas. Assim, ter uma crença envolveria a relação com um símbolo mental, sendo que a crença herdaria suas propriedades semânticas do símbolo mental que aparece na relação. Os processos mentais (pensar, perceber, aprender, etc.) envolvem interações causais entre os estados relacionais, tais como ter uma crença. As propriedades semânticas das palavras e sentenças que nós proferimos seriam herdadas, por sua vez, das propriedades semânticas dos estados mentais que a linguagem expressa.

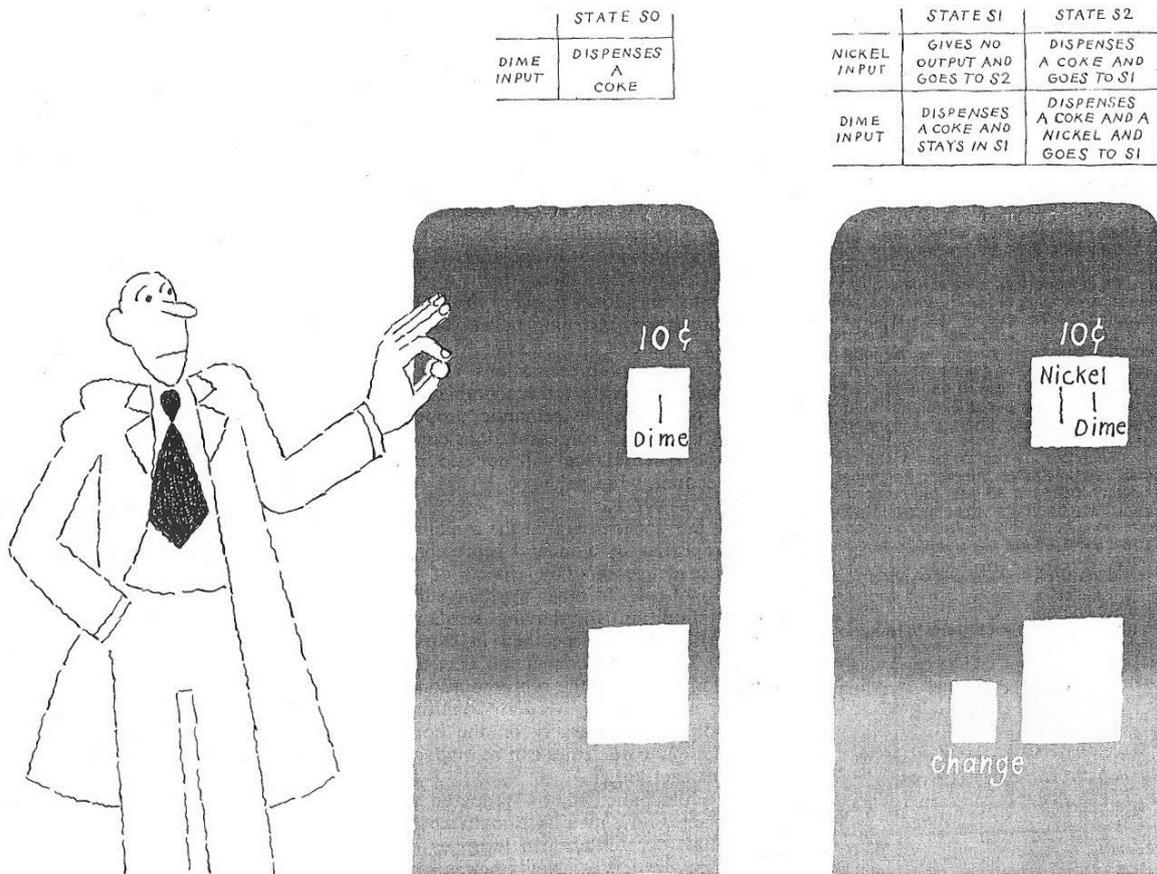
Associar as propriedades semânticas dos estados mentais com as dos símbolos mentais é totalmente compatível com a metáfora computacional porque é natural pensar o computador como um mecanismo que manipula símbolos. Uma computação é uma corrente causal de estados computacionais e os elos da corrente são operações sobre fórmulas semanticamente interpretadas, de acordo com um código de máquina. Pensar

em um sistema (como o sistema nervoso) como um computador é levantar questões sobre a natureza do código em que ele computa e sobre as propriedades semânticas dos símbolos do código. Na verdade, a analogia entre mentes e computadores de fato implica a postulação de símbolos mentais. Não existe computação sem representação.

A abordagem representacional da mente, entretanto, antecede consideravelmente a invenção do computador. Ela é um ancestral da epistemologia clássica, que é uma tradição que inclui filósofos tão distintos como John Locke, David Hume, George Berkeley, René Descartes, Immanuel Kant, John Stuart Mill e William James.

Hume, por exemplo, desenvolveu uma teoria representacional da mente, que incluía cinco pontos. Primeiro, existem “Ideias”, que são uma espécie de símbolo mental. Segundo, ter uma crença envolve acalentar uma ideia. Terceiro, os processos mentais são associações causais de Ideias. Quarto, as Ideias são como imagens. E quinto, as Ideias possuem suas propriedades semânticas em virtude de sua semelhança com aquilo que representam: a Ideia de John refere-se a John porque ela se parece com ele.

Os psicólogos cognitivos contemporâneos não aceitam os detalhes da teoria de Hume, embora mantenham muito de sua essência. As teorias da computação oferecem uma explicação muito mais rica dos processos mentais do que a mera associação de Ideias. E somente alguns poucos psicólogos ainda



**DUAS MÁQUINAS DE COCA-COLA** ilustram a diferença entre o behaviorismo (a doutrina de que não há causas mentais) e o mentalismo (a doutrina de que há causas mentais). Ambas as máquinas fornecem uma coca-cola por 10 cents e têm estados que são definidos por referência ao seu papel causal. A máquina à esquerda é behaviorista: seu único estado (*S0*) é definido somente em termos da entrada (*input*) e da saída (*output*). A máquina da direita é mentalista: seus dois estados (*S1*, *S2*) devem ser definidos não somente em termos da entrada e da saída, mas também em termos um do outro. Colocando de outra forma, a saída da máquina de coca-cola depende tanto do estado em que a máquina está quanto da entrada. O filósofo funcionalista defende que estados mentais são interdefinidos, como os estados internos da máquina de coca-cola mentalista [*dime* = 10 cents, *nickel* = 5 cents, *change* = troco].

pensam que a imagética é o principal veículo da representação mental. Entretanto, a ruptura mais significativa com a teoria de Hume é o abandono da semelhança como explicação das propriedades semânticas das representações mentais.

Muitos filósofos, a começar por Berkeley, têm argumentado que há algo muito errado com a sugestão de que a relação semântica entre um pensamento e aquilo a que o pensamento se refere poderia ser uma relação de semelhança. Considere o pensamento de que John é alto. É claro que o pensamento é verdadeiro apenas no que diz respeito ao fato de John ser alto. Uma teoria das propriedades semânticas de um pensamento deveria, portanto, explicar como esse pensamento específico está relacionado a esse fato específico. De acordo com a teoria da semelhança, o processo do pensamento envolve a posse de uma imagem mental que mostra John sendo alto. Em outras palavras, a relação entre o pensamento de que John é alto e o fato de ele ser alto é como a relação entre um homem alto e seu retrato.

A dificuldade com a teoria da semelhança é que qualquer retrato que mostre John sendo alto deve mostrar também um monte de outras coisas a seu respeito: se ele está vestido ou nu; se ele está deitado, em pé ou sentado; se ele possui uma cabeça ou não; e assim por diante. Um retrato de um homem alto que está sentado assemelha-se tanto a um homem sentado quanto a um homem alto. Na teoria da semelhança, não está claro o que é que distingue os pensamentos sobre a altura de John dos pensamentos sobre sua postura.

A teoria da semelhança parece deparar-se com paradoxos a todo momento. A possibilidade de se construírem crenças envolvendo relações com representações mentais semanticamente interpretadas depende claramente de se ter uma explicação aceitável da origem das propriedades semânticas das representações mentais. Se não é a semelhança que irá oferecer essa explicação, qual é a alternativa?

A ideia atual é de que as propriedades semânticas de uma representação mental são determinadas por aspectos de seu papel funcional. Em outras palavras, uma condição suficiente para ter propriedades semânticas pode ser especificada em termos causais. Essa é a conexão entre o funcionalismo e a teoria representacional da mente. A psicologia cognitiva moderna espera fortemente que essas duas doutrinas possam se sustentar reciprocamente.

Nenhum filósofo está preparado para dizer exatamente como o papel funcional de uma representação mental determina suas propriedades semânticas. Apesar disso, o funcionalista reconhece três tipos de relação causal entre estados psicológicos envolvendo representações mentais, que podem servir para estabelecer as propriedades semânticas das representações mentais. Os três tipos são relações causais entre estados mentais e estímulos, entre estados mentais e respostas, entre os próprios estados mentais.

Considere a crença de que John é alto. Presume-se que os seguintes fatos, que correspondem respectivamente aos três tipos de relação causal, são relevantes na determinação das propriedades semânticas da representação mental envolvida na crença. Primeiramente, a crença é um efeito normal de certas estimulações, tais como ver John em circunstâncias que revelam sua altura. Em segundo lugar, a crença é a causa normal de certos efeitos comportamentais, como afirmar que “John é alto”. Em terceiro lugar, a crença é uma causa normal de outras crenças e também um efeito normal de outras crenças. Por exemplo, qualquer pessoa que acredite que John é alto, muito provavelmente acreditará também que alguém é alto. Ter a primeira crença é causalmente suficiente para ter a segunda. E qualquer pessoa que acredite que todos numa sala são altos e também que John está nessa sala, provavelmente acreditará que John é alto. A terceira crença é um efeito normal das duas primeiras. Resumindo, o funcionalista sustenta que a proposição expressa por uma certa representação mental

depende das propriedades causais dos estados mentais em que as representações mentais aparecem.

A concepção de que as propriedades semânticas das representações mentais são determinadas por aspectos de seu papel funcional ocupa um lugar central nos trabalhos atuais dentro das ciências cognitivas. Entretanto, essa concepção pode não ser verdadeira. Muitos filósofos que não têm simpatia pela virada cognitivista na psicologia moderna duvidam de sua veracidade, e muitos psicólogos provavelmente a rejeitariam com base nessa apresentação breve e rudimentar que esbocei. No entanto, mesmo em sua forma rudimentar, muita coisa pode ser dita a seu favor. Ela legitima a noção de representação mental, que vem se tornando cada vez mais importante para a teorização em todos os ramos das ciências cognitivas. Os recentes avanços na formulação e no teste de hipóteses sobre a

natureza das representações mentais, em áreas que vão desde a fonética até a visão computacional, sugerem que o conceito de representação mental é fundamental para as teorias empíricas da mente.

O behaviorista tem rejeitado o apelo à representação mental porque ela contraria sua visão dos mecanismos explicativos que podem constar nas teorias psicológicas. Entretanto, a ciência da representação mental está florescendo. A história da ciência mostra que quando uma teoria bem sucedida entra em conflito com um princípio metodológico, é geralmente o princípio que é abandonado. Assim, o funcionalista tem flexibilizado as restrições behavioristas sobre as explicações psicológicas. Provavelmente, não há maneira melhor de se decidir o que é metodologicamente permissível em ciência do que investigar o que uma ciência bem sucedida requer.

p. 132 (b)

## **BIBLIOGRAFIA** (p. 148)

*Leitores interessados em mais explicações sobre os assuntos cobertos pelo artigo podem achar útil a seguinte lista de publicações.*

BLOCK, NED (org.) (1980), *Readings in Philosophy of Psychology*, vol. 1, Harvard University Press. [Leituras em filosofia da psicologia]

FODOR, JERRY A. (1968), *Psychological Explanations: An Introduction to the Philosophy of Psychology*, Random House Inc. [Explicações psicológicas: uma introdução à filosofia da psicologia]

FODOR, JERRY A. (1975), *The Language of Thought*, Thomas Y. Crowell Co. [A linguagem do pensamento]

PUTNAM, HILARY (1975), *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge University Press. [Mente, linguagem e realidade: artigos filosóficos]

VENDLER, ZENO (1972), *Res Cogitans: An Essay in Rational Psychology*, Cornell University Press. [Res cogitans: um ensaio em psicologia racional]