

Filosofia da Mente

Samuel C. Bellini-Leite

Compreender a mentalidade é atualmente uma das dificuldades filosóficas e científicas mais polêmicas, com pouco consenso envolvido, e principalmente, poucas propostas satisfatórias. Os principais problemas incluem o de dizer como a mente pode ter poder causal na matéria (causação mental – veja Amaral, 2001; KIM, 1998), o problema de determinar a individualidade de uma pessoa portadora de uma mente (problema da identidade pessoal – veja Coelho, 2020), o problema de determinar o conteúdo que a mente carrega (problema do conteúdo mental – veja De Menezes, 2020.), e como esse conteúdo porta um significado que faz referência ao mundo (problema do significado e da intencionalidade – veja Leclerc, 2018), dentre outros. Podemos dizer que todos esses estão de certa forma relacionados a um problema geral, conhecido como o problema mente-corpo (veja Fodor, 1981). A resposta para esse problema deve explicar a natureza da mente, e, havendo diferença em relação ao corpo, como esses dois se relacionam.

O problema da mente perpassa toda a história da filosofia e envolve diversas áreas de estudo, como a epistemologia, a metafísica, a lógica, a linguagem e a ética. Sendo assim, é comum encontrar em qualquer filósofo, teses que hoje podemos relacionar ao campo da filosofia da mente. Um filósofo historicamente lembrado pelo início de questionamentos especificamente mentais é o francês René Descartes (1641). Entretanto, é comum apontar o início formal da filosofia da mente ao início do século XX, a partir de interesses da filosofia analítica em compreender a relação de conceitos e proposições com o mundo (RUSSELL, 1921). Inclusive, esses dois pontos estão relacionados. A filosofia da mente, propriamente dita, surgiu de críticas à René Descartes feita por filósofos analíticos do século XX que começavam a se interessar pelo tópico (RYLE, 1949).

1. Críticas ao Dualismo de Substância

A proposta desenvolvida por Descartes é conhecida hoje como dualismo de substância, ou ao menos, consideram este um dos principais responsáveis pela elaboração e defesa dessa forma específica de dualismo (MASLIN, 2007). Essa visão de mente pode ser compreendida como parte do que Gilbert Ryle (1949) chamou de “doutrina oficial”. O dualismo de substância propõe que a mente é uma substância diferente do corpo, que possui características intrinsecamente distintas. Em particular, enquanto o corpo tem extensão, ocupando espaço físico e podendo ser dividido em pedaços, Descartes argumentou que a mente, ou a alma, é uma substância imaterial e sem extensão. Portanto, esta não ocuparia espaço físico e nem poderia ser dividida, característica que garantiria inclusive a imortalidade da alma após a morte do corpo. Se a alma é uma substância diferente do corpo que possui características distintas e que não podem ser divididas, então, pelo raciocínio, ela poderia continuar íntegra após o desmanche do corpo em pedaços (veja Descartes, 1641; e Maslin, 2007).

Evidentemente, a proposta de Descartes (1641) é mais complexa e elaborada do que esse parágrafo descrito acima, e o leitor fica convidado a se dirigir ao trabalho original. Entretanto, a filosofia da mente, propriamente dita, não se inicia em Descartes, mas na crítica ao dualismo de substância. Passaremos agora, então, para a explicação de críticas ao dualismo de substância, que convencem filósofos, desde o início do século XX, que outra resposta precisa ser dada ao problema mente-corpo.

O primeiro problema é que o dualismo de substância desconsidera a biologia ao tratar a mente como um fenômeno “tudo-ou-nada”. É intuitivo, em nossa cultura, pensarmos em seres humanos com almas, mas e se fôssemos pensar em animais com almas? Inicialmente não parece má ideia, por exemplo, um cachorro ter alma, se nós também temos. Entretanto, se a gente estende esse conceito na linha evolutiva, ele começa a ficar bastante estranho. Se a gente supuser que os animais possuem almas, teríamos que dizer que formigas, peixes, lesmas, estrelas do mar, dentre outros seres do reino animal, possuem uma substância imaterial que de alguma forma acopla aos seus corpos durante a vida, mas perduram de uma forma não-física após a morte.

Descartes, inclusive, considerava animais como meras máquinas que

reagem ao mundo de uma forma quase que automática, ou seja, eles não possuiriam mentalidade, seriam corpos vazios de almas, mais próximos aos robôs (ou autômatos) do que seres humanos. Entretanto, atualmente, as evidências da biologia comparativa, do estudo de células e tecidos, da neurociência comparativa, e etologia comportamental, mostram que animais possuem uma consciência similar à nossa em vários aspectos (EDELMAN, BAARS e SETH, 2005; EDELMAN e SETH, 2009). Recentemente, um grupo de neurocientistas e estudiosos da consciência se reuniram para assinar uma declaração sobre a consciência de animais, a partir dessas evidências (LOW et al., 2012). Logo, esta visão dos animais enquanto autômatos foi cientificamente contestada. Assim, qualquer mentalidade que aplicamos aos seres humanos, precisamos aplicar também, pelo menos em parte, à diversos outros animais.

Entretanto, podemos, é claro, traçar um corte evolutivo e supor, por exemplo, que apenas mamíferos possuem alma. Mas assim não estaríamos coerentes com a noção de mudança gradativa da biologia evolutiva. Isto é, as mudanças na história da biologia ocorreram de forma gradativa durante a evolução e a mentalidade precisa seguir também este princípio. Além disso, o conceito começa a perder o sentido; poderíamos nos questionar como seria uma alma de uma ratazana, ou se morcegos teriam almas. O conceito de alma parece ter uma conotação moral, de individualidade e justiça. Por exemplo, culturalmente tendem a dizer que as ações dos seres humanos durante a vida influenciam após a morte de uma forma moral; se alguém for justo, ele terá benefícios, caso contrário terá prejuízos. É difícil pensar em como um morcego entraria nessa história moral. Por diversas razões, limitar a alma para os mamíferos ainda assim seria um corte bastante contraintuitivo.

Outro problema para o dualismo de substância é explicar a causação mental. Enquanto é possível concordar que explicar a causação mental é uma dificuldade para qualquer tese que proponha falar de mente, este problema é especialmente fatal para o dualismo de substância. Quando um teórico assume que existem estados mentais, é necessário que ele seja capaz de explicar como os estados mentais influenciam o corpo. Por exemplo, quando eu *penso* (estado mental: pensar) em mover meu braço e *decido* (estado mental: decidir) agir, como podem estes estados gerarem um movimento físico no meu corpo? Bem, pelo dualismo de substância, o pensamento é responsabilidade da alma, uma substância imaterial, sem extensão, não-física. Em contrapartida, uma interação

causal depende de mudanças físicas. Quando movimentamos uma peça de xadrez, estamos usando nossas mãos para aplicar força física ao objeto e deslocá-lo para a posição desejada. É difícil compreender como a alma, responsável pelo desejo e intenção, poderia desencadear correntes elétricas e químicas no cérebro que levam este comando para o braço e para as mãos.

De uma forma geral, é difícil de entender como haveria comunicação entre a alma e as partes do cérebro (como o córtex motor) que as evidências neurocientíficas mostram estar relacionadas ao movimento e à intenção do movimento. Descartes atribuiu à glândula pineal, uma estrutura pequena e única no cérebro, a responsabilidade de interação com a alma. Entretanto, além de não estar relacionada às funções de alto-nível como o pensamento (usualmente atribuídas ao córtex pré-frontal), simplesmente apontar uma estrutura física de comunicação não resolve o problema. Ainda precisaríamos saber como a alma (não-física) interage com a glândula pineal (física). Esse problema é especialmente fatal para o dualismo de substância pois, mesmo se supormos algum fenômeno que gera essa interação, uma força vinda de uma estrutura não-física mudaria todo o fechamento causal da física. Para que haja uma alma capaz de causar, precisamos, em essência, abandonar o que a física nos mostrou sobre a como funciona a estrutura do universo (mas veja Melo e Pessoa, 2016). Em termos morais e epistemológicos, uma pessoa que acredita em alma precisa dizer ao físico que ela compreende a estrutura causal do universo de uma forma que o físico não compreende, o que certamente não parece provável. Precisamos, na filosofia, ter respeito com as ciências e com os especialistas, não podemos formar preceitos que vão contra teses bem construídas da ciência, sem termos bons motivos para isso. No mínimo, enquanto podemos dizer que a filosofia fornece uma liberdade de explorar possibilidades que outras áreas acadêmicas geralmente não fornecem, podemos também dizer que quanto mais esses preceitos se distanciam da ciência, mais eles perdem a relevância para a comunidade acadêmica.

O fato de a intenção do movimento poder ser mapeada ao córtex, e não apenas o início causal da corrente elétrica relacionada ao movimento (HAGGARD, 2005), nos lembra de mais um problema para o dualismo de substância: se as funções como intenção, personalidade, memória, identidade pessoal, pensamento, desejo, podem ser funções do cérebro, o que restaria para ser função da alma? O que o dualismo de substância nos traz é um choque de

funções da alma e do cérebro, e se torna difícil determinar o que seria responsabilidade de cada substância.

Há um caso, clássico na história da neurociência, de um homem que teve o cérebro perfurado em um acidente de trabalho, mas sobreviveu, Phineas Gage (veja Damasio, 1996). Relatos da época dizem que houve uma mudança drástica em sua personalidade e seu caráter. Na realidade, como este é um caso antigo, existem diversos problemas relacionados à sua interpretação e à sua validade. De toda forma, atualmente existem inúmeros exemplos de alterações de personalidade, memória, identidade pessoal, assim como mudanças contraintuitivas na mentalidade das pessoas que sofrem lesões ou degenerações cerebrais, que mostram como essas funções necessariamente precisam estar relacionadas ao cérebro durante a vida (SACKS, 1985).

Diante dessa divisão de funções entre cerebrais e aquelas da alma, o teórico dualista precisa decidir, não se sabe como, o que é exatamente função de cada substância. Como essas funções são alteradas através de lesões e degenerações cerebrais, parece natural concluir que na realidade todas essas funções que antes eram atribuídas à alma, são simplesmente funções cerebrais. Afinal, o cérebro já sabemos que existe.

Considere a hipótese de que a alma levaria consigo, após o perecer do corpo, algumas marcas de uma pessoa que lhe garanta sua individualidade, como a memória, pensamento e personalidade. Agora vamos considerar um caso de uma pessoa que sofreu de doenças naturais na infância, bullying na adolescência e por isso mudou sua forma de ser. Finalmente, aos trinta anos começou a apresentar sintomas dissociativos, como perda de identidade e personalidades múltiplas. Por fim, sofreu de Alzheimer quando idoso, fazendo com que perdesse sua memória e alterasse ainda mais sua personalidade. Após a morte, o que essa alma levaria consigo? As características genéticas sem alterações da infância? A forma como ela viveu na adolescência? Quais das personalidades durante seus episódios dissociativos? E por fim, levaria suas memórias intactas ou apenas o que restou após o Alzheimer?

É muito raro que tudo isso aconteça a uma pessoa, apesar de haver muita comorbidade na psicopatologia, mas, de toda forma, todas essas dificuldades acontecem com certa frequência e com muitas pessoas. Assim, essas dificuldades psicopatológicas nos mostram que a nossa memória, nossa personalidade, e até mesmo nossa identidade estão intimamente ligadas a um

momento da nossa vida, a um estado do nosso cérebro e do nosso corpo. Não faz sentido dizer que esses estados continuariam após a morte, pois mesmo na vida tendemos a perder esses estados. Eles são estados, e assim, não são infinitos, são passageiros, como tudo na biologia.

Há razões para acreditar que o dualismo de substância é uma explicação homuncular. Um homúnculo é um homem pequeno que imaginamos para problemas que não conseguimos resolver. Quando não se compreendia sobre a estrutura do esperma, por exemplo, imaginava-se que poderia haver um homúnculo escondido que iria desenvolver na gestação. Além dessa explicação ser falsa, ela também não é efetiva. O problema teórico com essa explicação é que se o homúnculo estivesse no esperma, precisaríamos ainda de uma explicação sobre como o homúnculo é criado. Logo, para explicar a gestação, postulavam um homúnculo, mas a entidade postulada em si também não era explicada, fazendo com que a necessidade de explicação simplesmente mudasse de foco do esperma para uma entidade postulada, o homúnculo.

O mesmo pode ser dito sobre teses do dualismo de substância. Se a resposta para a pergunta “como nós pensamos?” for “temos uma alma que pensa”, temos como consequência uma terceira pergunta “como a alma pensa?”; isto é, nada foi explicado; a explicação apenas se deslocou do cérebro para a alma, e alma parece ser ainda mais difícil de explicar do que o cérebro.

Vale notar que há uma forma de explicação homuncular válida. Por exemplo, termos como “grupos neurais”, “sistemas” e “representações” podem agir de forma homuncular em explicações. Entretanto, nesses casos, as entidades postuladas são mais simples do que aquelas explicadas, podem existir e podem ajudar a conectar dois níveis de explicação, como o celular e o funcional, por exemplo. Esta forma de explicação, conhecida como funcionalismo homuncular, é justificável e ainda possui adeptos (LYCAN, 1990; SANT’ANNA, 2012). Isto é, em suma, a dificuldade da explicação homuncular do dualismo é essencialmente que a entidade postulada é mais complexa, misteriosa, ou seja, mais difícil de explicar, e implausível de existir, do que o fenômeno a ser explicado em si.

Por fim, há um último problema com uma característica não-física da alma, a falta de extensão. Pelo dualismo de substância, a alma não ocupa lugar físico no espaço, ela não pode ser maior ou menor, larga ou delgada, ela não pode ser mensurada. Entretanto, quando teóricos do dualismo de substância

afirmam que cada pessoa tem a sua alma, eles estão sendo incoerentes. Se alma não tem extensão, então não seria possível contá-la, não poderíamos dizer que “João tem uma alma” e “Letícia tem outra alma”, pois ao fazer isso estaríamos dando a ela extensão. Uma entidade sem extensão não poderia ser individualizada dessa forma e muito menos ser atribuída a uma pessoa como se fosse uma coisa única e separável, isto vai contra a própria definição de alma. Há aqui então, um problema considerado muito grave na filosofia, uma teoria contraditória. Isso faz ela não apenas ser falsa sobre o nosso mundo, mas impossível de ser verdadeira em qualquer mundo possível (veja Maslin, 2007).

O problema de contar almas é difícil de perceber inicialmente, mas em geral, é muito fácil gerar outros problemas para a doutrina oficial, basta tentar. Por exemplo, se considerarmos a possibilidade de espíritos desincorporados que falam, poderíamos nos questionar como eles poderiam produzir sons, se sons são ondas produzidas na laringe, pelas cordas vocais, devido à passagem de ar dos pulmões. Poderíamos questionar como eles supostamente encostariam em algo, e ao mesmo tempo passariam por paredes. Enfim, se uma pessoa estiver disposta a se questionar, ela consegue se livrar de mitos culturais da doutrina oficial com facilidade (PENNYCOOK, et al. 2012). Tente você também desenvolver novas objeções.

Abandonar o dualismo de substância pode ser particularmente difícil para algumas pessoas, pois ele geralmente está associado às crenças religiosas, que podem ser preciosas para estas. Entretanto, muitas vezes, e isto é muito comum mesmo, para pensar de uma forma filosófica precisamos estar dispostos a largar e reconstruir nossas crenças mais profundas, que mais se associam com quem pensamos ser, antes de começar a estudar. Esse tipo de mudança também pode ser a mais valiosa para nossa vida intelectual, é essa abertura que permite que um filósofo avalie as teses por si, considerando as razões para elas serem verdadeiras ou boas teses, sem que elas tenham vínculo com nossa personalidade. Afinal nossa personalidade é individual, já teses sobre o mundo devem valer igualmente para todos, afinal o mundo é o mesmo.

O trabalho da filosofia da mente se inicia quando o filósofo percebe que o dualismo de substância está equivocado e se pergunta então como ele pode ser substituído. Isto é, quando se elimina o dualismo de substância, o problema não está resolvido, pelo contrário, a partir dessa dissolução é que o filósofo da mente precisa entrar para responder os problemas de como a mente funciona no

mundo.

2. Propostas Anti-Mentalistas

A primeira postura dos filósofos analíticos devido à derrota do dualismo de substância foi a de igualar o mentalismo ao dualismo. Se todo tipo de mentalismo for dualismo, então o mentalismo, por consequência, também estaria equivocado. Podemos chamar de mentalista, qualquer tese em filosofia da mente ou psicologia que utilize termos mentais em sua explicação de base, isto é, termos como acreditar, conhecer, compreender, raciocinar, sonhar, imaginar, tentar, querer, pretender, dentro outros. Podemos chamar de anti-mentalista qualquer tese que tenha uma proposta para se livrar desses termos, ou que tente evitar falar de estados mentais. Existem diversas maneiras que filósofos propuseram se livrarem desses termos: os teóricos das teses de identidade (SMART, 1959; PLACE, 1970), os eliminativistas (DENNETT, 1991a; CHURCHLAND, 1984; LEAL-TOLEDO, 2018), os ilusionistas (FRANKISH, 2016) e os behavioristas (RYLE, 1949; SKINNER, 1957).

Teóricos de identidade, como Smart (1959) e Place (1970), não acreditam ser possível eliminar o vocabulário de estados mentais em teses da filosofia da mente, mas se perguntam, será possível falar de eventos internos e processos privados, sem assumir uma posição dualista na qual sensações e imagens mentais formam uma categoria separada dos processos físicos, com os quais são correlacionados? A estratégia que traçam é de argumentar que esses estados são idênticos à estados cerebrais. Isto é, um estado mental, como sonhar, seria idêntico ao processo cerebral relacionado a sonhar, não haveria diferença a ser apontada.

Place (1970) não acredita que seja possível derrubar sua tese em termos lógicos, afirma que tentativas de demonstrar logicamente a falsidade da afirmação “a consciência é um processo no cérebro” falham na diferenciação de tipos de significado da palavra “é”. O “é” de definição caracteriza uma junção de sujeito e predicado que torna a afirmação verdadeira por definição: um quadrado é equilátero; vermelho é uma cor; entender uma instrução é ser capaz de agir apropriadamente nas circunstâncias apropriadas. Já em afirmações que utilizam o “é” de composição, o sujeito e o predicado da frase não possuem uma relação de necessidade, e a veracidade precisa ser verificada empiricamente: sua mesa

é uma caixa velha, o seu chapéu é um conjunto de palha amarrado por um fio, uma nuvem é uma massa de pequenas partículas em suspensão no ar. Para Place (1970) a afirmação “a consciência é um processo no cérebro” está usando o “é” de composição, o que faz com que sua hipótese seja uma afirmação empírica. Ele acredita que seria preciso investigar empiricamente para descobrir se sua teoria da identidade é verdadeira ou falsa.

A teoria da identidade tem ainda duas vertentes, a versão tipo-tipo e a versão token-token. Na versão token-token, cada evento cerebral único é idêntico a um estado mental único, não haveria como formar categorias de identidade. Dessa forma, quando eu penso em um gato, há um evento cerebral ocorrendo que é idêntico a esse meu pensamento de gato, e apenas esse evento aqui e agora forma essa identidade. Por outro lado, na versão tipo-tipo, poderíamos dizer que ver é idêntico a ter certos estados do lobo occipital do córtex, em diversas pessoas. A identidade seria igual em diversos cérebros, sendo possível formar categorias. Esta versão supera a teoria token-token pois por ela seria possível ter um base para formar uma ciência, pois existe a possibilidade de encontrar verdades generalizáveis para diversas pessoas. Enquanto na teoria de token-token, se cada estado mental específico corresponde a um estado cerebral específico de um dado momento, não seria possível ter uma ciência que faça um estudo da ligação de estados de identidade, pois eles seriam únicos a cada ocorrência. Por outro lado, parece simplesmente falso que, quando dois sujeitos pensam em gato, os mesmos estados neurais estariam ativados. Assim, a teoria de identidade tipo-tipo é empiricamente falsa, pelo menos para diversas funções cerebrais que não possuem correspondências entre indivíduos.

Um problema geral dessa vertente de identidade é que em nosso melhor sentido do termo identidade, um conceito é idêntico a outro quando ele carrega todas suas propriedades. Mas um elefante rosa em nossa imaginação possui propriedades que não se aplicam aos neurônios que o produzem, os neurônios não são rosas. Assim, parece aceitável dizer que neurônios produzem a imaginação, mas não parece ser coerente dizer que nossa imaginação é idêntica aos neurônios, pois possuem propriedades distintas.

A corrente eliminativista (CHURCHLAND, 1984) desistiu de tentar mapear estados mentais em estados cerebrais, seja por tipo-tipo ou token-token. O que eles propõem é uma completa eliminação do uso de termos mentais em

qualquer explicação séria sobre a mente. Em vez de falarmos de dores, após estudos neurocientíficos descobrirem os estados cerebrais relacionados à dor, digamos a “Fibra C”, por mais que em nosso diálogo cotidiano usemos termos de estados mentais, a ciência pelo menos, deveria falar apenas dos estados cerebrais. Por exemplo, o cientista nãoalaria “ele está sentindo dor”, mas diria “estão ocorrendo fortes ativações de sua Fibra C”. Os eliminativistas acreditam que estados mentais são uma invenção da psicologia popular, a teoria cultural dos estados mentais. Eliminativistas variam sobre a utilidade da psicologia popular. Alguns acreditam que a psicologia popular apenas atrapalha em qualquer comunicação humana. Outros acreditam que ela foi útil evolutivamente e continua sendo útil culturalmente. Mas todos concordam que para a ciência, especificamente, esses termos precisam ser abandonados, como outros termos anticientíficos já foram no passado, como o *flogisto*, e evidentemente, a alma.

Uma versão atual das teses eliminativistas é o ilusionismo (Frankish, 2017). Esta tese difere termos mentais de estados qualitativos da consciência, como *qualia* (veja Nagel, 1974). Enquanto ela propõe que podemos falar de estados mentais como acreditar, pensar e desejar, ela não aceita a realidade de estados fenomenológicos; acredita que a introspecção é uma ilusão, apenas mais um viés cognitivo que sofremos. Entretanto, como ela não foca necessariamente em eliminar estados mentais básicos, ela não traz o conflito que as demais teses anti-mentalistas trazem para o atual capítulo.

Estas teses mencionadas acima ainda possuem um grande papel na filosofia da mente, mas podemos considerar o behaviorismo a maior corrente, em termos históricos, das teses anti-mentalistas. O behaviorismo ocorreu não apenas na filosofia com Bertrand Russell (1921), ou Ryle (1949), mas também na psicologia com Watson (1913) e Skinner (1957). Alguns teóricos behavioristas acreditam não haver uma mente para ser explicada, outros acreditam nos estados mentais, mas pensam que metodologicamente precisam ser ignorados para que haja uma ciência possível, e outros afirmam que conseguiremos compreender a mente melhor pelo comportamento do que pela introspecção. Apesar de haver diferenças entre todas essas propostas mencionadas, podemos dizer que todas são simpáticas, ao que Jerry Fodor (crítico do behaviorismo) chama de a tese essencial do behaviorismo.

Fodor (1968, p.51) chama de behaviorista qualquer doutrina que aceite a proposição ‘P’, onde a proposição P diz que para cada predicado mental que

pode ser colocado em uma explicação psicológica, precisa haver pelo menos uma descrição do comportamento com a qual ele tem uma conexão. Em contrapartida, qualquer doutrina que nega P pode ser chamada de mentalista. Fodor, diferentemente dos anti-mentalistas, diferencia o mentalismo do dualismo, onde o dualismo é qualquer doutrina em que mentes são compostas de uma substância irreduzível ao corpo. O dualista precisa ser um mentalista, mas um mentalista não precisa ser um dualista. Para o mentalista, basta afirmar que predicados mentais e predicados sobre o comportamento são logicamente distintos e ainda assim se aplicam sobre um mesmo e único corpo. Fodor (1968) argumenta que teóricos do início do século XX, quando discutindo entre o behaviorismo e o dualismo não perceberam que ambos poderiam estar errados. Fodor (1968) assegura ainda que enquanto argumentos de eliminação funcionam contra a ideia de múltiplas substâncias, eles não funcionam contra o mentalismo.

O behaviorista moderado não tenta eliminar os predicados mentais, ele meramente julga que esses predicados precisam ter um correspondente no comportamento. Fodor (1968) explica que um dos argumentos adotados pelo behaviorista é que nós conseguimos compreender os outros pelo comportamento, e o mentalismo pode nos levar à privação do conhecimento dos estados mentais, ou seja, à tese de que não podemos compreender os outros pelo comportamento. Como parece claro que sabemos quando alguém sente dor ou está feliz, então o mentalismo estaria nos levando para um problema que não condiz com a forma como compreendemos os outros.

Entretanto, Fodor (1968) argumenta que a ligação entre predicados mentais e predicados comportamentais não é necessária, mas apenas contingente, e não anômala (veja Fodor, 1989), por seguir regras *ceteris paribus*, ou seja, regras que 'geralmente ocorrem se tudo estiver dentro do esperado'. Isto é, o tipo de regra entre comportamentos e predicados mentais que temos possuem a seguinte forma "se alguém está gritando, batendo o pé, e socando uma parede ele estará, *ceteris paribus*, nervoso". Mas essa conexão *ceteris paribus* entre os predicados comportamentais e os predicados mentais não garante a tese P que o behaviorista precisa. Pelo contrário, essa conexão fraca parece garantir algo muito mais próximo ao mentalismo, pois haverá casos onde alguém grita, bate o pé e soca e mesmo assim não está nervoso, inclusive a pessoa pode estar nervosa e exibir sorrisos. Lembrando que, o mentalista

evidentemente concorda que muitas vezes conseguimos deduzir os estados mentais nos baseando no comportamento.

O psicólogo, que, com razão, preocupa-se em estudar a mente de forma científica, aplica o behaviorismo como uma ferramenta metodológica para supostamente compreender os fatos observáveis de uma forma neutra e eficaz. Para contrapor a isso, Fodor (1968) considera um caso em que a resposta B é condicionada ao estímulo A e uma resposta C ao evento B. É de se esperar que a apresentação de A elicite C mesmo que o par A-C nunca tenha sido condicionado. A tendência é supor que B tem um papel ocorrente mediador e não observável. Fodor (1968) considera plausível que B seja produzido de uma forma não observável e influencie a resposta C. Poderíamos ainda supor essa ocorrência com mais segurança se o par A-C gastar mais tempo para ocorrer do que outro par semelhante que não pareça ter um fator interno mediador. Fodor (1968) argumenta que se pesquisas mostrarem que a melhor hipótese para explicar o comportamento depende de eventos internos mediadores, sem um comportamento observável correspondente, o psicólogo teria que escolher entre abandonar a metodologia original ou abandonar a melhor explicação dos fatos. Considerando que nunca devemos abandonar a melhor forma de compreender os fatos, e que a metodologia inicial foi proposta justamente para cumprir o objetivo de explicar os fatos, parece que a solução mais aceitável é de abandonar o princípio metodológico.

Além das críticas específicas mencionadas para cada forma de anti-mentalismo, há uma crítica que se aplica, de forma fatal, a todas as teses anti-mentalistas, a permanência das proposições de identidade.

Mesmo supondo alguma forma de identidade (tipo-tipo ou token-token) entre estados mentais e físicos, as proposições que versam sobre essa identidade seriam contingentes e passíveis de análise. Então, para essa identidade existir, ainda assim, precisamos concordar que existem essas proposições de identidade que não podem ser formuladas sem o uso da linguagem mentalista. Por exemplo, a proposição “crenças são estados neurais X” inclui um termo mentalista, a saber: crenças. Se a ciência começar a usar o termo “estados X” em vez de crença, isto não muda o fato de haver essa ligação inicial baseada na proposição de identidade, que contém o termo crença e que, então, faz a referência a um estado mental. Em contraste, você pode eliminar, por exemplo, o termo “alma” da linguagem de identidade e não haverá perda

explicativa para a ciência.

O eliminativista pode acreditar se safar dessa crítica. Entretanto, estudando a metodologia da neurociência de forma séria, percebe-se que não há como fugir. Para ser possível um neurocientista atribuir uma função a uma região do cérebro, ele precisa pensar sobre os estados mentais, e conversar sobre os estados mentais com o sujeito da pesquisa. Para estudar a imaginação, ele fala para o sujeito do seu experimento, por exemplo “imagine um elefante rosa”. Então, nesse caso, não apenas ele usou seus próprios estados mentais para definir seu objeto de estudo em termos mentais – a imaginação, como atribuiu estados mentais ao sujeito do experimento utilizando termos mentais na comunicação, para apenas depois dizer que certa área do cérebro “Y” corresponde à imaginação. Posteriormente, dizer que não devemos falar de imaginação, e apenas devemos falar de Y, não tem valor algum real de eliminação, pois tudo que estaríamos fazendo é atribuindo um código Y (uma área cerebral específica, um neurotransmissor, uma reação química) para falar do estado mental imaginação, estaríamos, em essência, apenas mudando o nome da palavra, mas a referência continua sendo o estado mental. Conseguiríamos compreender o significado de “Y” pois sabemos, pela introspecção e pela psicologia popular o que é imaginação. Logo, até mesmo em teses eliminativistas, a crítica da permanência das proposições de identidade de Fodor (1958) se aplica de forma fatal.

Vale notar que apesar dos comentários acima, teses anti-mentalistas ainda formam uma corrente comum e aceita na filosofia da mente (veja Leal-Toledo, 2018).

3. Mentalismo de Processamento de Informação

As críticas que o behaviorismo sofreu (CHOMSKY, 1959; FODOR, 1968) abriram espaço para novas formas de falar sobre a mente. Como os predicados mentais precisam necessariamente existir na psicologia, o dilema disponível não é entre usar ou não usar predicados mentais, mas sim entre ter uma linguagem popular dos estados mentais (ideias, intenções, crenças, desejos) ou uma linguagem científica destes. A única corrente em toda a história da filosofia e da psicologia que forneceu uma linguagem científica para tratar dos estados mentais foi a ciência cognitiva, através da hipótese do processamento de

informação. Essa hipótese trouxe o mentalismo para destaque no estudo da mente.

A hipótese do processamento de informação diz que estados mentais são processos informacionais ocorrendo no cérebro humano, de forma semelhante a como um computador (material) roda um software (padrões de energia). Não se trata apenas de uma analogia. Apesar de nenhum filósofo acreditar que o cérebro é como um computador clássico inventado no século XX, uma grande parte concorda (FODOR, 1975; DENNETT, 1991a; CHALMERS, 1995) com a afirmação metafísica de que a mente é um processador de informação.

A principal mudança tecnológica que influenciou o início dessa vertente foi o surgimento de computadores, em especial as aplicações da inteligência artificial. Não apenas em termos de aplicações, mas também em termos filosóficos. Alan Turing, um matemático que auxiliou a decodificar códigos nazistas na segunda guerra mundial, propôs uma das primeiras formulações matemáticas de um computador (TURING, 1936) e também influenciou os estudos da mente através de questionamentos sobre como o pensamento está relacionado ao processamento de informação de computadores (TURING, 1950).

Uma máquina de Turing (1936) é uma abstração que, de forma geral, consiste em uma fita (infinita) dividida em quadrados onde um símbolo pode ser escrito; um leitor que considera apenas um quadrado de cada vez; e um conjunto de regras que funciona como uma tabela de instruções para o que a máquina faz. O leitor examina o símbolo na fita e prossegue conforme a tabela de instruções determina. A máquina, a qualquer momento, pode substituir algum símbolo na fita, mover o leitor para a esquerda ou para a direita ou parar. A partir da leitura de instruções e manipulação de símbolos, em princípio, essa máquina pode praticar qualquer tipo de computação. Esta máquina é uma abstração matemática. No entanto, a arquitetura Von Neumann foi planejada para instanciar fisicamente uma versão dessa máquina.

A arquitetura de Von Neumann (Von Neumann, 1945) é um projeto para implementar um computador digital. Consiste em uma unidade de processamento, uma unidade de controle, armazenamento primário e mecanismos de entrada/saída. A unidade de processamento contém uma unidade aritmética/lógica que executa instruções aplicando procedimentos aritméticos ou lógicos em um código binário e um registro que armazena

rapidamente etapas de tais operações. Este sistema é análogo ao que lê e aplica as alterações na fita no Máquina de Turing. Uma unidade de controle contém um registro de instruções, que por sua vez contém as instruções a serem processadas em um determinado momento, mediando o fluxo de processamento, fornecendo sinais de tempo e controle. É isso que torna a ideia de uma tabela de instruções fisicamente possível. Tais instruções são mantidas no armazenamento primário, juntamente com os dados que foram trabalhados na unidade de processamento. Portanto, trata-se de uma memória que é análoga à fita na Máquina de Turing, mas que também armazena as instruções da tabela de instruções. Mecanismos de entrada e saída são o que permitem comunicação entre este núcleo central com todas as suas outras partes, como seu disco rígido externo.

Bem, podemos ver que a implementação de um computador digital tem uma estreita semelhança arquitetônica com uma Máquina de Turing, e isso não seria verdade para os casos de um computador quântico, um computador funcionando em paralelo ou um cérebro. De toda forma, os primeiros proponentes da hipótese de processamento de informação não possuíam ligação com a neurociência e acreditavam ser possível falar de uma mente, de certa forma, independentemente de sua implementação neural.

Newell e Simon (1976) e Newell (1980) propuseram uma primeira base para se pensar como agentes dotados de um sistema de processamento de informação poderiam resolver problemas no mundo real, elaborando a hipótese do sistema de símbolos físicos.

Newell e Simon (1976, p.116) afirmam que “um sistema de símbolos físicos consiste em um conjunto de entidades, os símbolos, que são padrões físicos (por exemplo, elétricos) que podem ocorrer como componentes de outro tipo de entidade, chamada de expressão (ou estrutura de símbolo).” Operações manipulam tais símbolos, criando, copiando, editando e destruindo expressões. Ao contrário da Máquina de Turing, para um sistema de símbolos físico é essencial que este esteja em um mundo dos objetos. Isso porque um dos papéis fundamentais dos símbolos e expressões é o de designar objetos. Newell e Simon (1976, p.116) explicam que “uma expressão designa um objeto se, dada a expressão, o sistema pode afetar o próprio objeto ou se comportar de maneiras dependente do objeto” e, assim, ganhar “acesso” ao objeto, via expressão. Como Newell (1980, p.156) coloca, “ter X (o símbolo) é equivalente a ter Y (a coisa

designada) para a finalidade do processo P". Qualquer objeto pode ser designado por meio de símbolos e expressões. Portanto, Newell (1980) afirma que os números em uma máquina de Turing não são símbolos como no caso de sistemas de símbolos físicos. Isso porque, enquanto os dígitos de uma máquina de Turing funcionam como *tokens* de expressões, elas não são, no entanto, acopladas a alguma estrutura de dados.

Ter símbolos capazes de designar discretamente concede ao sistema algumas propriedades, como transitividade. Se X é um símbolo que designa Y e Y designa Z, então X designa Z. Agora, só obtemos isso se X receber um valor fixo e discreto, assim, não pode haver incerteza sobre o valor de um símbolo, como se teria sobre uma variável. Em outras palavras, um símbolo designa um único objeto. Essas propriedades se estendem para produtividade e composicionalidade (veja Fodor e Pylyshyn, 1988).

Se uma expressão designar um processo do sistema e se o sistema puder executar tais processos, diz-se que o sistema pode interpretar a expressão. Outra maneira de entender isso é que expressões têm algum tipo de significado para a máquina quando ações (ou processos) podem ser derivadas delas. Tal processo é ativado pelo controle (uma parte do sistema de símbolos). O que o controle faz é reunir, em um determinado momento, dados e operadores. Interpretar expressões e aplicar operadores é o que o sistema de símbolos físicos faz para a produção de novas expressões. Nas palavras de Newell (1980, p.159) "esta organização implica um requisito para a memória de trabalho no controle para gravar os símbolos para o operador [...]" e também que "[...] a memória de trabalho é uma característica invariante dos sistemas de símbolos". Vale notar que um dos constructos da psicologia cognitiva com maior base de evidência é a memória de trabalho (BADDELEY e HITCH, 1974) que possui essas características bastante semelhantes às aquelas mencionadas por Newell (1980).

O sistema por si só não necessariamente possui habilidades para resolver problemas. Assim, Newell e Simon (1976, p.120) explicam a resolução de problemas através da hipótese de busca heurística: "as soluções para problemas são representadas como estruturas de símbolos. Um sistema de símbolos físicos exercita sua inteligência no problema, resolvendo pela busca - isto é, gerando e modificando progressivamente as estruturas simbólicas até produzir uma estrutura de solução."

Como Newell e Simon (1976) explicam, para esses sistemas, os problemas surgem de uma forma específica. Para que haja um problema, deve haver um espaço de problema ou um espaço de estados. Esse espaço inclui todos os passos possíveis que o processo pode dar ao resolver o problema, incluindo o estado inicial, várias soluções e etapas intermediárias. Também deve haver um teste que indica o estado que a máquina deve estar para atingir um objetivo. E, finalmente, deve haver geradores de movimento, que são processos de modificação de uma situação no espaço de problema. Newell e Simon (1976, p.121) explicam que “a tarefa que um sistema de símbolos enfrenta, então, quando é apresentado a um problema e um espaço de problema, é usar seus recursos de processamento limitados para gerar possíveis soluções, uma após a outra, até encontrar uma que satisfaça o teste de definição de problemas”. O sistema precisa, entretanto, se organizar para que as soluções tenham uma alta probabilidade de ocorrerem. Um sistema de símbolos exibiria inteligência para um dado problema na medida em que mostra que tem alta probabilidade de encontrar soluções para o problema, em vez de caminhar aleatoriamente ou por tentativa e erro em um espaço de estados.

Assim, no caso de sistemas de símbolos físicos, um espaço de problema contém todas as formas de expressão que poderiam ser dadas à forma inicial considerando os geradores de movimento disponíveis. Assim, cada expressão obtida no processo é uma modificação da expressão inicial. Além disso, os geradores de movimento incorporam regras, como a lógica, aritmética ou outra. Estas regras contêm diretrizes gerais de resolução de problemas e outras especificidades. Por exemplo, uma regra geral seria algo como: “não produza contradições”. Uma específica estaria relacionada ao problema em questão, como “crie a menor expressão possível”. Em qualquer etapa, as modificações não são caóticas, mas sempre dependem dessas regras incorporadas pelos geradores de movimento e o estado atual do processo, um dado local no espaço do problema.

Pense na resolução de problemas de um sistema cognitivo como um caminhar sobre as diversas possibilidades de resposta a um problema. Os passos tomados em um espaço de problema formam ramificações de uma árvore. Se todos os estados (ou a maioria dos estados) de um espaço de problema foram visitados (pesquisa de força bruta), o processo incorreria em explosão computacional (não haveria tempo para ocorrer na realidade), a menos

que o espaço do problema fosse muito pequeno (o que não ocorre em problemas relevantes para o mundo). Assim, a busca deve ser heurística no sentido de que deve tomar o menor número de etapas possíveis para obter soluções.

Como Newell e Simon (1976, p.123-124) afirmam: “Nossa análise de inteligência igualou-a à capacidade de extrair e usar informações sobre a estrutura do espaço do problema, de modo a permitir que uma solução de problema seja gerada rápida e diretamente quanto possível” e “a tarefa da inteligência, então, é evitar a ameaça sempre presente da explosão exponencial de busca.”

Fodor (1975) propôs uma tese bastante popular de que nosso cérebro instancia algo como um sistema de símbolos físico de Newell (1980), o que ele chamou de linguagem do pensamento. Fodor (1975) argumenta ainda que essa linguagem é mais básica que a linguagem natural (português ou inglês) e que seria impossível aprender qualquer língua sem antes ter a linguagem do pensamento. Esses desenvolvimentos formam o que hoje é chamado de teoria computacional (ou representacional) da mente (FODOR, 1985). Interessantemente, essa proposta insere os estados mentais como algo natural do cérebro enquanto processador de informações, há um realismo dos estados mentais. Os estados mentais seriam, assim, *padrões*¹²⁵ elétricos, ou químicos, que formam símbolos que designam objetos do mundo real, e simulam espaços de problemas para serem investigados por buscas heurísticas ou qualquer outra técnica computacional/informacional de resolução de problemas. O mentalismo, ganha com essa proposta, sua primeira hipótese ontológica ligada ao fisicalismo ou materialismo na história. Em outras palavras, a mente é inserida em uma ontologia física (ou informacional), sem a necessidade de substâncias postuladas, afinal computadores existem, e sem a necessidade de teses eliminativistas.

Historicamente, a principal crítica à hipótese do processamento de informação é a analogia do quarto chinês. Searle (1980) gera um experimento de pensamento no qual uma pessoa em um quarto recebe a tarefa de traduzir

¹²⁵ A palavra “padrões” está em itálico para enfatizar que a ontologia da mente nessa proposta não é elétrica ou química, a mente é compreendida como padrões de atividade de processos físicos.

chinês com base em regras procedimentais e uma lista de instruções que pode acessar. A pessoa recebe letras em chinês, segue as instruções de modificação e passa a resposta adiante. Searle (1980) argumenta que essa pessoa não saberia falar chinês, ela apenas estaria aplicando regras de uma forma cega, assim como qualquer processador de informação. Dessa forma, ele desenvolve o argumento de que processadores de informação não possuem acesso ao significado do conteúdo que processam, e como seres humanos possuem tal acesso, eles não poderiam ser processadores de informação.

Apesar da crítica do quarto chinês ser repetida diversas vezes na história da filosofia da mente, assim como na cultura popular, interessantemente, é pouco discutido como Daniel Dennett (1980; 1991b) apresenta respostas bastante convincentes a esse problema. Dennett explica como a metáfora do quarto chinês está equivocada em colocar o homem responsável pelas manipulações como o agente que sabe o significado da linguagem, isso faria da teoria do processamento de informação uma proposta homuncular no sentido ruim. Na realidade, o significado da linguagem emergiria, para a metáfora estar correta, de diversos quartos desses em interação (DENNETT, 1980), cada um com seu funcionamento específico (sistemas cognitivos variados), todos interligados e mapeados pela evolução, por seleção natural, para reagir de certa forma e significar coisas específicas (DENNETT, 1991b). Apesar de funcionar na base da sintaxe, agentes cognitivos (máquinas treinadas ou cérebros) transformam um certo símbolo em outro apenas quando estão em certa relação semântica (FODOR, 1985; mas veja FODOR e PYLYSHYN, 2015).

Há, entretanto, um problema similar ao problema do significado, como repetido por Searle, que afeta de forma mais séria a hipótese do processamento de informação, o problema do *frame*, ou o problema da relevância (DREYFUS, 1972; PYLYSHYN, 1987). Este problema se refere à dificuldade encontrada por um sistema que trabalha com expressões representando o mundo para agir de forma fluida. As tentativas iniciais da inteligência artificial representacionista falharam, pois os robôs não conseguiam agir, por terem de considerar e catalogar inúmeras mudanças irrelevantes que acontecem no mundo. Este problema atualmente vem sendo tratado através de propostas incorporadas (CLARK, 1997; BROENS e GONZALEZ, 2006.) e através de representações probabilísticas e preditivas (CLARK, 2016; BELLINI-LEITE, 2018). Ainda, em termos técnicos, houve avanço considerável, que relativizou o problema, por exemplo, com o

desenvolvimento de *machine learning*, especialmente o *deep learning*, que fazem máquinas realizarem atualmente funcionalidades que antes eram pensadas serem impossíveis para máquinas em geral, por causa do problema do *frame*. Isso não significa que o problema tenha sido superado, mas não se trata mais de um problema fatal para a hipótese do processamento de informação, considerando que mudanças de perspectivas filosóficas e tecnológicas têm avançado em sua solução.

Em contrapartida, um problema fundamental não abarcado pela hipótese de processamento de informação é o problema dos sentimentos e da experiência. David Chalmers (1995) é um filósofo contemporâneo que aceita o mentalismo de processamento de informação, mas acredita que a mente, mais especificamente a consciência, não está completamente abarcada por essa tese. Enquanto a mente compreende estados inconscientes, como o plano motor, que permite que andemos de bicicleta sem prestar atenção nos detalhes do pedal, os quais poderiam ser explicados por processamento de informação, há também uma parte da mente, a consciência, que possui estados qualitativos, uma experiência de estar consciente (NAGEL, 1974).

O problema da experiência como elaborado por Chalmers (1995) diz que existem dois problemas da consciência, o problema fácil e o problema difícil. O problema fácil da consciência é aquele que pode ser respondido em termos de comportamento, fisiologia e teorias da cognição. Já o problema difícil, o problema da experiência, está em explicar como e por qual motivo uma experiência ocorre em conjunto com o comportamento e o processamento de informação. Da forma como a mente é descrita em termos científicos, por exemplo, pela neurociência cognitiva, não haveria motivo algum para a cognição vir acompanhada de experiência como dores, cores, sentimentos dentre outras. O problema pode ainda ser especificado em dois: (1) qual é a natureza da experiência (NAGEL, 1974); isto seria algo como encontrar uma forma de compreender onde o mundo fenomenológico se encontra em relação ao mundo objetivo; e, (2) ao explicar a natureza da experiência, deve-se ter também uma história sobre como essa experiência emerge do funcionamento cerebral, ou em geral, qual é a ligação da experiência com o comportamento e processamento de informação em seres vivos.

Chalmers (1996) tentou resolver o problema da experiência a colocando como algo fundamental na física, como são matéria e espaço, e também

cogitando o pampsiquismo, a tese de que todos os objetos possuem algum tipo de mentalidade. Entretanto, Chalmers (1995, 1996) foi reconhecido mais pela forma de caracterizar o problema do que pela resposta, pois sua resposta não trouxe consenso algum para a comunidade acadêmica.

Na realidade, existem inúmeras tentativas de resolver o problema da experiência como colocado por Chalmers (1995), aceitando o mentalismo e também acreditando haver a necessidade de maiores explicações sobre a natureza da experiência consciente. Podemos mencionar, por exemplo, o dualismo de propriedades e o monismo não redutivo com suas diferentes formas (veja Maslin, 2007), o monismo reflexivo (VELMANS, 2008), o monismo de triplo-aspecto (PEREIRA JR., 2015) e o naturalismo biológico (SEARLE, 1990).

Entretanto, ainda que essas propostas se mostrem preocupadas em explicar como pode haver diferentes níveis e como eles interagem, enquanto não houver uma explicação aceita de qual seria a natureza da experiência consciente, que não simplesmente redescreva a fenomenologia (BELLINI-LEITE, 2015), essas parecem meramente variações linguísticas; termos diferentes para ideias muito parecidas – e, enquanto proliferarem teorias, nenhuma será suficientemente consensual.

A verdade é que alguma explicação extra seria necessária para o mentalismo de processamento de informação resolver os problemas de Chalmers. Pense o caso dos sentimentos e das emoções. Não parece nem um pouco claro que robôs, que processam informação, terão sentimentos ou emoções, a não ser que adotemos posições behavioristas, eliminativistas ou ilusionistas sobre esses estados. Se aceitarmos a experiência como real, e a descrição fenomenológica como algo que ainda depende de mais explicações metafísicas, como acredita Chalmers (1996), então, intuitivamente, não temos motivo algum para acreditar que nossos computadores, se ficarem mais complexos, irão começar a ter sentimentos ou emoções. Esses aspectos parecem não serem apenas de processamento de informação, mas estarem ligados ao nosso corpo biológico, como argumenta Searle (1990) corretamente. Intuitivamente parece haver algo de especial na biologia, que gera estados de experiência em diversos seres vivos, o que não é o caso na robótica. Não parece ser plausível dizer que uma câmera de supermercado consegue ver, da mesma forma que um cachorro, nem mesmo se essa câmera estiver ligada a um sistema de inteligência artificial que processa a imagem com facilidade, pois, ainda assim

não haveria para a câmara a experiência de enxergar.

Logo, apesar de o mentalismo de processamento de informação ser a corrente melhor aceita na neurociência cognitiva (GAZZANIGA, IVRY e MANGUN, 2006) e estar ainda muito forte na filosofia da mente (DAVENPORT, 2012; COLTHEART, 2012), esta hipótese parece ainda em desenvolvimento. Ela ainda se desenvolve em relação à natureza das representações mentais, enquanto paralelas (CLARK, 1989), probabilísticas e preditivas (CLARK, 2016), incorporadas (CLARK, 1997), estendidas (CLARK e CHALMERS, 1998; COELHO, 2017), simbólicas (FODOR e PYLYSHYN 1988), além das propostas enativistas (CARVALHO, 2018; DI PAOLO e THOMPSON, 2014), na esperança de conseguir resolver os problema do significado (FODOR E PYLYSHYN, 2015), relevância (PHYLYSHYN, 1987) e abdução (FODOR, 2001; VITTI-RODRIGUES, MATULOVIC e GONZALEZ, 2017), e precisa de acréscimos à teoria que deem conta da experiência consciente (CHALMERS, 2004), para que não haja a necessidade de novas propostas eliminativistas ou behavioristas para esconder a nossa lacuna de conhecimento.

4. Conclusão

O leitor pode pensar que a história da filosofia da mente foi de alguma forma circular e que, como teses anti-mentalistas e mentalistas também são problemáticas, o dualismo de substância seria novamente um concorrente apto. Isto está equivocado, enquanto o dualismo é nobre no sentido de tentar responder questões da consciência, em vez de negá-las, o problema é que estas respostas estão fundamentalmente equivocadas. Isto é, enquanto o anti-mentalismo e o mentalismo contemporâneo possuem problemas a serem resolvidos – alguma versão dos dois pode estar correta¹²⁶ –, o dualismo de substância é o único que já foi suficiente e convincentemente demonstrado estar equivocado, tanto por argumentos quanto por evidências científicas. Portanto a negação do dualismo de substância continua sendo o ponto de partida da

¹²⁶ Por exemplo, a adoção do processamento de informação para explicar termos da psicologia popular como crenças e desejos, associado ao ilusionismo (FRANKISH, 2017) para responder ao problema da experiência, ou associado à uma proposta da virtualidade da experiência (REVONSUO, 1995) são caminhos bastante plausíveis.

filosofia da mente. O que passa a ser necessário é o desenvolvimento de novas teses que resistam aos problemas que ainda faltam ser enfrentados. Não é exagero afirmar que uma possível resolução desses problemas afetaria em grande parte diversas outras áreas da filosofia, como a epistemologia, a linguagem, a lógica, a ética e a metafísica. Este é um ponto de trabalho fortíssimo atualmente, que precisa ainda de muita pesquisa e apoio. Assim, o leitor está convidado a estudar o campo para além dessa introdução, para que possa no futuro desenvolver teses que contemplem as evidências, ao mesmo tempo sendo capaz de atingir os problemas descritos no capítulo, assim como outros que podem ser encontrados na literatura da filosofia da mente.

Questões de Revisão

1. O que é o dualismo de substância?
2. Relate como o dualismo de substância está ligado às crenças culturais.
3. Explique três críticas ao dualismo de substância.
4. Explique por qual motivo o dualismo de substância foi considerado superado mesmo considerando que todas as demais propostas também possuem problemas.
5. O que são termos mentais?
6. Qual é a diferença entre a teoria da identidade mente-cérebro e o eliminativismo?
7. Explique uma crítica de Fodor às teses anti-mentalistas.
8. Descreva uma dificuldade de igualar um estado mental à um comportamento.
9. Por que o mentalismo não precisa ser um dualismo de substâncias?
10. O que é um sistema de símbolos físicos de Newell?
11. Qual é a ontologia da mente na hipótese do processamento de informação?
12. O que ainda está em desenvolvimento na hipótese do processamento de informação?
13. Qual é o problema difícil da consciência de Chalmers?

Questões de Discussão

1. Sugira quais das críticas ao dualismo de substância você considera mais

eficaz.

2. Por que ideias de alma ainda são prevalentes em nossa cultura?
3. Como é a mente de um animal?
4. Animais possuem qualidades do sentido que jamais conseguiremos ter?
5. A sua mente é seu cérebro? Há diferença?
6. Imagine um elefante rosa. Qual é a ontologia desta imagem?
7. Como propriedades aparentemente não-físicas da mente podem interagir com o mundo físico-causal?
8. Acredita que robôs possam um dia sentir dor e prazer? Por quê?
9. Procure por exemplos de ilusão de cor na internet (como o vestido preto e azul ou branco e dourado; ou o tênis rosa ou verde) e responda, a cor é uma propriedade do objeto ou da mente humana? Existe alguma posição intermediária?
10. Qual é a ontologia de um desejo?
11. Tente definir o que são crenças.
12. Há ainda um caminho plausível para teses anti-mentalistas?
13. Quais são as semelhanças e as diferenças entre mentes e computadores?

Sugestão de leitura

A filosofia da mente (TEIXEIRA, 2016) é um tópico vasto que faz conexões com diversas ciências como a psicologia, inteligência artificial, linguística, biologia, neurociências, antropologia, física, dentre outras. Além disso, há uma pluralidade de visões e não há, muito claramente, uma ordem histórica de autores que deve ser seguida. O que pode ser feito, entretanto, é uma seleção de alguns problemas da filosofia da mente que o leitor se interessa para que busque conteúdo sobre esses problemas em diversos autores. Esse capítulo apresenta alguns desses problemas, mas não todos. Assim, seria interessante começar os estudos com um manual mais completo de introdução à filosofia da mente (MASLIN, 2007; LEAL-TOLEDO, GOUVEA e ALVES, 2018).

Tendo selecionado um problema, a sugestão é deixar o livro de introdução e obter conhecimentos mais específicos sobre o problema. Por exemplo, para complementar o atual capítulo seria interessante o leitor interessado ler os diálogos originais entre Searle (1980; 1990) e Dennett (1980; 1991b), para formar sua opinião sobre o quarto chinês (veja também Vargas,

2014, para uma comparação dos dois autores). O leitor pode ver as diferença no tratamento da consciência através de propostas eliminativistas (CHURCHLAND, 1984; DENNETT, 1991a; LEAL-TOLEDO, 2018) e propostas que consideram a experiência qualitativa (CHALMERS, 1996; CHALMERS, 2004). Pode também, através de um mesmo autor, ver mudanças em relação a noção de representação e o problema do *frame* (CLARK, 1989, 1997, 2016).

Alternativamente, é importante verificar teses fundamentais e gerais que marcaram a história do estudo da mente, como o behaviorismo (RYLE, 1949; SKINNER, 1957; WATSON, 1913) e a versão simbólica do processamento de informação (NEWELL, 1980; FODOR, 1985; FODOR e PYLYSHYN, 1988), e eventualmente chegar em sugestões atuais sobre como enfrentar o problema da consciência (METZINGER, 2004; VELMANS, 2008; PEREIRA JR., 2015; FRANKISH, 2017).

Por fim, vale ressaltar que o recorte feito por esse capítulo não representa a única forma de fazer filosofia da mente, existe também uma corrente ligada à tradição fenomenológica (ABATH e CAMINHA, 2012; CAMPOS, 2017; NOË, 2004), trabalhos mais fortemente ligados à filosofia analítica (MENDONÇA, 2007; LEWIS, 1972) e também aqueles baseados na filosofia da linguagem (BARWISE e PERRY, 1981; PERINI-SANTOS, 2014; SEARLE, 1969).

Referências

- ABATH, A; CAMINHA, I. Merleau-Ponty e o fisicalismo. *Revista de Filosofia Aurora*, v.24, n.35, p.615-638, 2012.
- AMARAL, F. Causação Mental: Onde Estivemos e Onde Estamos. *Psicologia: Teoria e Pesquisa*, Brasília, v.17, n.3, p.235-244, 2001.
- BADDELEY, A.; HITCH, G. Working Memory. *Psychology of Learning and Motivation*, v.8, p.47-89, 1974.
- BARWISE, J.; PERRY, J. Situations and attitudes. *The Journal of Philosophy*, v.78, n.11, p. 668-691, 1981.
- BROENS, M.C.; GONZALEZ, Maria Eunice Quilici. Um estudo do conhecimento não proposicional no contexto da teoria da cognição incorporada e situada. *Manuscrito (UNICAMP)*, v.9, p. 729-751, 2006.
- BELLINI-LEITE, S. Explicação metafísica, causação e interações neuro-astrocitárias: uma réplica a “o conceito de sentimento no monismo de triplo

- aspecto". *Kinesis-Revista de Estudos dos Pós-Graduandos em Filosofia*, v.7, n.14, 2015.
- BELLINI-LEITE, S. Surfing Uncertainty: predições situadas?. *Ciências & Cognição*, v.23, n.2, p.255-262, 2018.
- CAMPOS, V. Pondo A Mão Na Consciência (Incorporada): sobre o estatuto do corpo no pensamento de Jean-Paul Sartre. *ConTextura*, v.9, n.11, 2017.
- CARVALHO, L. A Realização da Mente: Crítica da Teoria Enativa ao Conceito de Propriedade Emergente. In: LEAL-TOLEDO, G. L.; GOUVEA, R. A. S.; ALVES, M. A. S. (Org.). *Debates Contemporâneos em Filosofia da Mente*. 1ed. São Paulo: FiloCzar, p. 113-129, 2018.
- CLARK, A; CHALMERS, D. The extended mind. *analysis*, v.58, n.1, p.7-19, 1998.
- CLARK, A. Connectionist minds. In: *Proceedings of the Aristotelian Society*. Aristotelian Society, Wiley, 1989. p. 83-102.
- CLARK, A. *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press, 1997.
- CLARK, A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, 2016.
- CHALMERS, D. Facing up to the problem of consciousness. *Journal of Consciousness Studies*. v.2, n.3, 1995.
- CHALMERS, D. *The conscious mind: In search of a fundamental theory*. Oxford university press, 1996.
- CHALMERS, D. Como Podemos Construir Uma Ciência Da Consciência? (ALVES, M.; RAMOS, J. Tradução). *Prometheus-Journal of Philosophy*, n. 33, (Trabalho original publicado em 2004), 2020.
- CHOMSKY, N. A Review of B. F. Skinner's Verbal Behavior. *Language*, v.35, n.1, p.26-58, 1959.
- CHURCHLAND, P. Matéria e consciência (Maria Clara Cescato, trad.). São Paulo: Editora Unesp, (Trabalho original publicado em 1984), 2004.
- COELHO, B. *A hipótese da mente estendida: um debate acerca dos limites da cognição*. (Dissertação - Unesp) Marília, 2017.
- COELHO, J. Mente consciente, cérebro e corpo: casos de gêmeos siameses. *PRINCIPIOS*, v. 27, p. 91-113, 2020.
- COLTHEART, M. *The cognitive level of explanation*. Australian Journal of Psychology, v. 64, n. 1, p. 11-18, 2012.

- DAMÁSIO, A. *O erro de Descartes: emoção, razão e o cérebro humano*. Editora Companhia das Letras, 1996.
- DAVENPORT, David. *Computationalism: Still the only game in town*. *Minds and machines*, v. 22, n.3, p.183-190, 2012.
- DE MENEZES, J. Sobre a viabilidade da noção de conteúdo mental estreito. *Princípios: Revista de Filosofia* (UFRN), v.27, n.53, p. 115-134, 2020.
- DENNETT, D. *Consciousness explained*. New York: Black Bay Books, 1991a.
- DENNETT, D. *Ways of establishing harmony*. In Brian P. McLaughlin (ed.), *Dretske and His Critics*. Blackwell, 1991b.
- DESCARTES, R. *Meditações sobre Filosofia Primeira* (F. Castilho, trad.). Campinas: Editora Unicamp, (Trabalho original publicado em 1641), 2004.
- DI PAOLO, E.; THOMPSON, E. *The enactive approach*. *The Routledge handbook of embodied cognition*, p. 68-78, 2014.
- DREYFUS, Hubert L. *O que os computadores não podem fazer: crítica da razão artificial*. A Casa do Livro Eldorado SA, (Trabalho original publicado em 1972), 1975.
- EDELMAN, D., BAARS, B.; SETH, A. Identifying hallmarks of consciousness in non-mammalian species. *Consciousness and Cognition*, v.14, p.169–187, 2005.
- EDELMAN, D.; SETH, A. Animal consciousness: a synthetic approach. *Trends in Neurosciences*, v.32, n.9, 2009.
- FODOR, J. *Psychological Explanation: an introduction to the philosophy of psychology*. Nova Iorque: Random House, 1968.
- FODOR, J. *The Language of Thought*. New York: Crowell, 1975.
- FODOR, F. A. O problema mente-corpo. TCFC3 – Filosofia das Ciências Naturais, São Paulo: FFLCH-USP, (Trabalho original publicado em 1981) 2011. Disponível em: <<http://opessoa.fflch.usp.br/sites/opessoa.fflch.usp.br/files/Fodor-Port-5.pdf>>. Acesso em: 16 de Julho de 2020.
- FODOR, J. Fodor's guide to mental representation: The intelligent auntie's vademecum. *Mind*, v.94, n.373, p. 76-100, 1985.
- FODOR, J. Making mind matter more. *Philosophical Topics*, v.17, n.1, p.59-79, 1989.
- FODOR, J. *The Mind Doesn't Work That Way*. Cambridge: MIT Press, 2001.
- FODOR, J; PYLYSHYN, Z. *Connectionism and Cognitive Architecture: A Critical*

- Analysis. *Cognition*, v.28, 1988.
- FODOR, J; PYLYSHYN, Z. *Minds without meanings: An essay on the content of concepts*. Mit Press, 2015.
- FRANKISH, K. *Illusionism: as a Theory of Consciousness*. Andrews UK Limited, 2017.
- GAZZANIGA, Michael S.; IVRY, Richard B.; MANGUN, George Ronald. *Neurociência cognitiva: a biologia da mente*. Artmed, 2006.
- HAGGARD, P. Conscious intention and motor cognition. *Trends in cognitive sciences*, v.9, n.6, p.290-295, 2005.
- KIM, J. *Mind in a physical world: An essay on the mind-body problem and mental causation*. MIT press, 1998.
- LEAL-TOLEDO, G. Ensaios de um fisicista estrito. In: LEAL-TOLEDO, G.; GOUVEA, R.; ALVES, M. *Debates Contemporâneos em Filosofia da Mente*. CEP, v.5852, p.95-112, 2018.
- LECLERC, A. *Uma introdução à filosofia da mente*. Appris Editora e Livraria Eireli-ME, 2018.
- LOW, P.; PANKSEPP, J.; REISS, D.; EDELMAN, D.; VAN SWINDEREN, B.; KOCK, C. *The Cambridge Declaration on Consciousness*. Cambridge, Francis Crick Memorial Conference on Consciousness in Human and non-Human Animals, 2012.
- LEWIS, D. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, v.50, n.3, p.249-258, 1972.
- LYCAN, W.G. *Mind and Cognition: A Reader*. Oxford: Blackwell Publishers, 1990.
- MASLIN, K. T. *Introdução à filosofia da mente* (FJR da Rocha, trad.). Porto Alegre: Artmed, (Trabalho original publicado em 2007), 2009.
- MCCARTHY, J.; HAYES, P. Some philosophical problems from the standpoint of artificial intelligence. In: B. Meltzer & D. Michie (Eds.), *Machine Intelligence 4*. Edinburgh, Scotland: Edinburg University Press, 1969.
- MELO, L.; PESSOA, O. O dualismo interacionista não precisa violar leis de conservação da física. In: MACHADO, Nivaldo; SOBRINHO, Elizeu de Oliveira Santos. *Filosofia da mente, ciência cognitiva e o pós-humano: para onde vamos?*, de Monica Aiub, Maria Eunice Quilici Gonzalez e Mariana Cláudia Broens (Organizadoras). Griot: Revista de Filosofia, v. 13, n. 1, p. 119-121, 2016.
- MENDONÇA, W. *Fisicismo Não-Reducionista: Uma atitude sem conteúdo*

- congnitivo? Sobre o desafio de Bas Van Fraassen. *Principia: an international journal of epistemology*, v.11, n.2, p. 171-186, 2007.
- METZINGER, T. *Being no one: The self-model theory of subjectivity*. MIT Press, 2004.
- NAGEL, Thomas. Como é ser um morcego? *Revista da Abordagem Gestáltica: Phenomenological Studies*, v.19, n.1, p. 109-115, (Trabalho original publicado em 1974), 2013.
- NEWELL, A. Physical symbol systems. *Cognitive Science* 4, 1980.
- NEWELL, A.; SIMON, H. Computer Science as Empirical Inquiry: Symbols and Search. *The ACM Communications*, v.19, n.3, 1976.
- NOË, Alva. *Action in perception*. MIT press, 2004.
- PENNYCOOK, G. Analytic cognitive style predicts religious and paranormal belief. *Cognition*, v.123, n.3, p.335-346, 2012.
- PERINI-SANTOS, E. Contextualismo. In: BRANQUINHO, João; SANTOS, Ricardo. *Compêndio em linha de problemas de filosofia analítica*, Lisboa: Centro de Filosofia da Universidade de Lisboa, 2014.
- PEREIRA JR., A. O Conceito de Sentimento no Monismo de Triplo Aspecto. *Kínesis, Edição Especial – Debate*, v.7, n.15, p.1-24, 2015.
- PLACE, Ullin T. Is consciousness a brain process?. In: *The mind-brain identity theory*. Palgrave, London, 1970. p. 42-51.
- PYLYSHYN, Z. (ed.) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. New Jersey: Ablex Publishing Corporation, 1987.
- REVONSUO, A. Consciousness, dreams and virtual realities. *Philosophical Psychology*.v.8, n.1, p.35-58, 1995.
- RYLE, G. *O conceito de espírito* (ML Nunes, Trad.). Lisboa: Moraes, 1949/1970.
- RUSSELL, B. *A análise da mente*. Zahar Editores, (Trabalho original publicado em 1921), 1974.
- SACKS, O. *O homem que confundiu sua mulher com um chapéu: e outras histórias clínicas*. Editora Companhia das Letras, (Trabalho original publicado em 1985), 2016.
- SANT'ANNA, A. Funcionalismo homuncular e o problema dos qualia. *Filogênese*. Marília, v.5, p.1-15, 2012.
- SEARLE, J. *Speech acts: An essay in the philosophy of language*. Cambridge university press, 1969.
- SEARLE, J. *Mentes, cérebros e programas*. In: *Cérebros, Máquinas*

- eConsciência. Org. João de Fernandes Teixeira, São Carlos: EDFUSCar, (Trabalho original publicado em 1980), 1996.
- SEARLE, J. *O mistério da consciência e discussões com Daniel C. Dennett e David J. Chalmers*. Paz e Terra, (Trabalho original publicado em 1990), 1998.
- SMART, J. Sensations and brain processes. *The Philosophical Review*, v.68, n.2, p.141-156, 1959.
- SKINNER, B. *Verbal behavior*. New York: Appleton Century-Crofts, 1957.
- TEIXEIRA, J. *O que é filosofia da mente*. 2ed, Porto Alegre: Editora Fi, 2016.
- TURING, A. *Computação e inteligência*. (FC Hanssem, Trad.) Em J. de F. Teixeira (Org.), *Cérebros Máquinas e Consciência: Uma Introdução à Ciência da Mente*. São Carlos: Editora da UFSCar, (Trabalho original publicado em 1950), 1996.
- TURING, A. On computable numbers with an application to the entscheidungs problem. *Proceedings of the London Mathematical Society*2, v.42, n.1, 1936.
- VARGAS, G. *A consciência na filosofia contemporânea da mente*. 1ª. ed. Novas Edições Acadêmicas, 2014. 164p.
- VELMANS, M. Reflexive monism. *Journal of Consciousness Studies*, v.15, n.2, p.5-50, 2008.
- VITTI-RODRIGUES, M.; MATULOVIC, M.; GONZALEZ, M. Informação-processo e abdução. In: ALVES, Marcos Antonio; MARTÍNEZ-ÁVILA, Daniel; GRACIO, Maria Cláudia Cabrini. *Informação, conhecimento e modelos*, p.131-151, 2017.
- VON NEUMANN, J. The First Draft Report on the EDVAC. *IEEE Annals of the History of Computing*, v.15, n.4, (Trabalho original publicado em 1945), 1999.
- WATSON, J. A psicologia como o behaviorista a vê. *Temas em psicologia*, v.16, n.2, p. 289-301, (Trabalho original publicado em 1913), 2008.