

Causação Mental: Onde Estivemos e Onde Estamos¹

Felipe Amaral²

City University of New York – CUNY

RESUMO: O presente trabalho tem como propósito situar o problema da causação mental dentro do debate contemporâneo em filosofia da mente. A partir dos trabalhos de autores como Smart, Putnam, Davidson e, principalmente Kim, delinearemos as mais importantes tentativas de solução do problema desde fins da década de cinquenta. Defenderemos que a teoria da identidade, o funcionalismo e o monismo anômalo não se mostraram alternativas viáveis para o amelhoramento de nosso entendimento de como propriedades, ou tipos, mentais podem ter poder causal em um mundo físico. Concluiremos, por fim, que nem mesmo a superveniência mente-corpo, assim como indicado nos mais recentes trabalhos de Kim, foi capaz de oferecer uma saída satisfatória ao problema, colocando-nos assim diante de um impasse: abraçamos um projeto reducionista ou aceitamos o epifenomenismo do mental. Não apresentaremos uma defesa do projeto reducionista, porém sugeriremos que este nos parece menos problemático do que a alternativa epifenomenista.

Palavras-chave: causação mental; teoria da identidade de tipo; funcionalismo; monismo anômalo.

Mental Causation: Where we Were and Where we Are

ABSTRACT: This paper aims at generally situating the issue of mental causation in the contemporary debate in the philosophy of mind. From the work of authors such as Smart, Putnam, Davidson and, more importantly, Kim, I shall try to delineate what came to be known, since the fifties, as the main efforts to solve the problem of mental causation. In consonance with Kim's work, I shall claim that the type-type identity theory, functionalism, and anomalous monism, in general, are not easily understood as viable positions to increase our understanding of how mental properties, or types, may be seen to have causal potency in a physical world. I shall conclude that since the relation of supervenience does not offer us a legitimate solution to the problem, as indicated by Kim's most recent work, we are left with two stunning options: mind-body reductionism or epiphenomenalism. I shall choose none, although I suggest reductionism to be less problematic.

Key-words: mental causation; type identity theory; functionalism; anomalous monism.

Comumente consideramos nossos estados mentais causas reais de nossos comportamentos. Por exemplo, ao nos perguntarem o porquê de termos atravessado uma rua movimentada fora da faixa para pedestres, podemos responder que assim agimos porque esta teria sido a única maneira de assegurarmos o nosso objetivo: chegar ao ponto de ônibus antes do ônibus partir. Por esta explicação evidencia-se a assunção de que nosso comportamento externo, ou periférico, foi causado por estados internos como desejos e crenças,

de modo que o ato de atravessar a rua esteve causalmente conectado ao desejo de alcançar o ônibus e à crença de que não atravessando fora da faixa não o alcançaríamos. Uma outra assunção comum em relação à nossa vida mental parece ser a de uma ou mais crenças serem causadas por outras como, por exemplo, a crença de que o saldo bancário está negativo em conjunção à crença de que o novo carro custou-nos muito dinheiro causam a crença de que estamos gastando mais do que devemos. Este exemplo mostra-nos um caso de causação mental-ao-mental, enquanto que o anterior apresenta-nos uma situação de causação mental-ao-físico. Alternativamente, seria possível também contemplarmos casos de causação físico-ao-mental, como o do contato de fótons com a retina causando a percepção de algo colorido.

Descartes formulara tais relações mente-corpo em termos de duas substâncias, a coisa pensante e a coisa extensa, cada qual com características essenciais distintas (viz., a consciência e a extensão, respectivamente). Por isso, o problema da causação mental põe-se para Descartes na forma da seguinte pergunta: como é possível substâncias essencialmente distintas interagirem causalmente? Descartes respondeu-nos que a mente relaciona-se com o corpo através de uma pequena porção de nosso cérebro, a glândula pineal. Esta glândula seria, em uma rede causal, o ligante entre a mente e o corpo. Não obstante isso—apontar-se-á— a glân-

1 Este trabalho descende de uma dissertação de graduação apresentada ao Departamento de Filosofia da Universidade de Brasília no primeiro semestre de 2001. Estiveram presentes na banca avaliadora os Professores Norberto Abreu e Paulo Abrantes. Agradeço enormemente ao Prof. Norberto por comentários em versões anteriores deste trabalho e pela cordialidade em convidar-me a preparar uma versão reduzida do texto original para a publicação nesta Revista. Agradeço-lhe também por sua paciência e incentivo, sem as quais eu não teria conseguido preparar este texto. Agradeço muitíssimo também ao Prof. Paulo Abrantes, orientador do trabalho do qual este descende, pelos inúmeros e pacientes comentários, sugestões e correções em versões preliminares. Sinto-me muito privilegiado por ter tido a oportunidade de trabalhar com o Professor Paulo. Responsabilizo-me inteiramente por eventuais erros e imprecisões que porventura permanecem aqui.

2 Endereço: 983 50th Street Apt. 2F – Brooklyn, NY – 11219 – felipeamaral@earthlink.net

dula pineal também é parte do mundo físico, tendo assim as mesmas características fundamentais das coisas físicas. Deste modo, como seria possível haver interação causal entre a glândula pineal e a mente, dado serem essas substâncias essencialmente heterogêneas?

O problema cartesiano da interação causal entre mente e corpo não é mais o nosso problema da causação mente-corpo. Isso porque a discussão do problema mente-corpo é apresentada, correntemente, dentro dos limites do fisicalismo.³ Desta maneira, o nosso problema da causação mental não é apresentado em termos de um dualismo de substâncias, mas sim como um problema gerado a partir de um dualismo de propriedades (Kim, 1998). Sob esta diferença jaz a posição metafísica de que não existem coisas não-físicas, isto é, a realidade é exaurida por entes físicos. Ademais, tal pressuposição metafísica, chamada fisicalismo de substâncias, parece-nos favorável do ponto de vista metodológico, pois se a aceitamos, podemos pesquisar o que há através de conceitos e procedimentos característicos das ciências naturais e isto bastar-nos-ia para uma abrangente história explicativa do mundo físico. Poder-se-ia naturalmente disputar essa postura metafísica e a sua correspondente metodologia; todavia, não é nosso intuito aqui discutir essa problemática nem muito menos resolvê-la. Acreditamos que o fisicalismo de substâncias seja o ponto de partida da nossa discussão e por isso mesmo simplesmente suporemos esses limites metafísicos no que se segue.

Apesar de não padecermos do problema cartesiano da causação mental, padecemos de um problema similar. Mesmo supondo uma realidade exaurida por entes físicos, não solucionamos por isso o problema concernente à natureza da interação entre mentalidade e corpo.⁴ Isso porque ao fisicalismo de substâncias é comum que se una uma visão bifurcada das características, ou propriedades, das substâncias físicas, deslocando-se assim o tradicional dualismo mente-corpo ao âmbito das propriedades, ou características, das coisas. Nesse sentido, alguns entes exemplificariam, ou realizariam, tanto propriedades físicas quanto propriedades mentais; poder-se-ia falar, por exemplo, em um ente com a propriedade ‘... está no estado mental de tipo dor’ ou a propriedade ‘... está no estado neural de tipo ativação da fibra-C’. Destarte, o *nosso* problema mente-corpo mostra a sua face quando nos perguntamos, por exemplo, sobre a relação entre a dor que sentimos ao encostarmos em um pedaço de brasa e o fenômeno neurobiológico que ocasiona aquela sensação; as dificuldades surgem, sobretudo, quando examinamos as cadeias causais que relacionam o estado mental, nesse exemplo a dor, o estado neural e os outros estados men-

tais e físicos envolvidos. Portanto, a pergunta a se fazer é: qual o papel da dor na rede causal que a liga ao nosso movimento de escape, dado que o estado neurobiológico que a ocasiona também têm influência causal em nosso comportamento? De maneira mais geral, a pergunta crucial parece ser a seguinte: dado que para cada estado mental há um estado físico correspondente que o ocasiona, qual o papel causal do primeiro face à influência causal óbvia do segundo na ocorrência de outros estados físicos (ou mentais)?

Essas perguntas, com efeito, configuram o núcleo do problema mente-corpo. No que se segue, seguindo Kim (e.g., 1993, 1997, 1998) em sua transparente discussão dessa problemática, apresentaremos as três grandes tentativas contemporâneas em tratar a possibilidade de estados mentais terem poder causal em um mundo físico. A primeira tentativa, exemplificada em Smart (1959), propôs resolver o problema mente-corpo dissolvendo-o, isto é, reduzindo a mente ao corpo. Conseqüentemente, o problema especial da causação mental dissolver-se-ia naturalmente (apesar deste último não ter sido, aparentemente, o ímpeto de Smart ao ter tratado da natureza das sensações). Defenderemos que a assim chamada ‘identidade de propriedades’, ou ‘identidade de tipo’,⁵ falhou em seu projeto maior por não ter respondido satisfatoriamente às críticas que lhe foram dirigidas (e.g., Putnam, 1967; Kripke, 1980). A segunda grande tentativa de solução do problema, apresentada e defendida por Donald Davidson (1970), também não nos parece inteiramente satisfatória. Davidson, a partir de uma posição monista em relação à natureza dos eventos (viz., todos os eventos são físicos), confere aos eventos mentais poder causal em um mundo físico fechado causalmente, sem no entanto salvar um papel causal atuante para as propriedades, ou tipos, mentais, dos quais os eventos mentais são ocorrências; nesse sentido, a posição davidsoniana deixa de abranger um ponto bastante importante do problema da causação mental. A terceira tentativa, possivelmente a que mais adeptos angariou nas últimas décadas, o funcionalismo, pelo menos em algumas de suas formulações também não nos parece uma solução viável do problema da causação mental. Apesar de algumas vantagens face à teoria da identidade de tipo, a posição funcionalista deixa de especificar, assim como no caso da posição davidsoniana, a diferença causal que as propriedades mentais fazem em um mundo físico. Por conseguinte, concluiremos, que de modo geral essas três tentativas falham em nos apresentar uma imagem plausível da metafísica da mente, que faça justiça às nossas aspirações fisicalistas sem ignorar a realidade aparentemente diversa dos fenômenos mentais por nós experienciados.

Em um último momento de nosso texto analisaremos a noção de superveniência mente-corpo. Defenderemos que em última análise a superveniência mente-corpo não só não resolve o problema da causação mental como impõe um dilema que nos deixa poucos caminhos viáveis para a sua so-

3 Apesar do termo ‘fisicismo’ ser mais adequado à formação de palavras da gramática portuguesa, privilegiamos o uso do termo ‘fisicalismo’ porque queremos nos divorciar de conotações de uma defesa de uma *linguagem* da física como privilegiada e redutora de todas as outras. Além disso, lembramos que o termo ‘fisicalismo’ é abonado pelo novo Aurélio. [v. Aurélio séc. XXI]

4 Usamos aqui o termo ‘mentalidade’ porque preferimos evitar as conotações reificantes da expressão ‘ter mente’, no sentido de se ter uma *coisa* não-física

5 Para os propósitos deste texto não diferenciaremos propriedades e tipos.

lução (Kim, 1998). Se nossa argumentação for correta, a visão de que tipos de estados mentais têm existência robusta, independente dos tipos de estados físicos que os ocasionam, se torna dificultosa. Nos restam a redução ou o epifenomenismo (Kim, 1998).

Smart e a Identidade Mente-Corpo

J.J.C. Smart, em fins da década de cinquenta, publicou um trabalho que veio a se tornar um clássico da literatura sobre a natureza das sensações. Nesse trabalho, Smart defendeu a tese de que sensações não passam de processos cerebrais. Em termos mais precisos, tipos de sensações são tipos de ativações neurais. Nessa asseveração está contida a extensão da posição smartiana, pois para ele não é o bastante que ocorrências de estados mentais sejam ocorrências de estados físicos.

A consciência de tal diferença entre tipos e ocorrências é muito importante para que percebamos com clareza a abrangência das identidades em questão. Dizer que toda ocorrência de um estado mental é uma ocorrência de um estado físico é identificar exemplificações delimitadas temporalmente. No organismo O, no instante de tempo t, a sensação X é a ativação de uma fibra neural. Já a identificação de tipos de estados mentais a tipos de estados físicos não tem indexação temporal alguma. Simplesmente, ao aceitarmos as identidades de tipo, segue-se que ter um estado é ter outro, sempre. Assim, dizer que dor é ativação da fibra-C, no sentido de identificação de tipo, é dizer que *todas* as ocorrências de dor são, ao fim e ao cabo, ocorrências de ativação da fibra-C. Portanto, a diferença essencial entre identidade de ocorrência e identidade de tipo é que a primeira estabelece uma relação puramente *de facto* entre mente e corpo, enquanto que a segunda estabelece uma relação mais forte, sendo assim mais estável e portanto permitindo-nos fazer previsões e dar explicações. Se lhe for perguntado o porquê de sempre que há dor, há ativação da fibra-C, o adepto da identidade de tipo poderá dizer 'porque dor é ativação da fibra-C'. Já o adepto da identidade de ocorrência não pode oferecer nenhuma explicação mais geral para tal relação entre dor e ativação neural; a única coisa que o último pode afirmar é que naquela situação específica a dor não foi nada mais nada menos do que uma ativação neural.

Não obstante isso, ao tratarmos as teses de Smart devemos prosseguir com bastante cuidado, pois, apesar dele ter defendido a identificação entre tipos de estados mentais e neurais, ele não asseverou a necessidade dessas identidades. Com efeito, Smart expressamente afirmou que as identidades das quais ele falava deveriam ser entendidas como contingentes. Portanto, assim como os cientistas naturais descobriram que 'trovoada = raio' ou que 'calor = agitação molecular', descobrir-se-ia oportunamente que tipos de sensações seriam nada mais nada menos que tipos de ativações nervosas.⁶

O caráter contingente das identidades mente-corpo por Smart defendidas parece ter sido conseqüência de uma visão não muito clara das relações entre os conceitos de aprioricidade, necessidade e analiticidade. Segundo Smart, a palavra 'dor' não significa 'ativação da fibra-C', pois podemos falar da dor que sentimos sem que nós ou nossos interlocutores saibam nada de neurofisiologia. Além disso, a determinação do valor de verdade dessas identificações mente-corpo depende em última análise da pesquisa empírica. Para que saibamos se 'dor = ativação da fibra-C' é uma sentença verdadeira, precisamos testar essa hipótese; simplesmente usando a razão não podemos descobrir o seu valor de verdade. Portanto, as identidades smartianas não são nem analíticas nem a priori. Segue-se daí, consoante Smart, que essas identidades são contingentes, no sentido em que valem em nossa situação, ou mundo possível, mas não em outros. Isto é, como 'dor' não significa o mesmo que 'ativação da fibra-C' e 'dor = ativação da fibra-C' pode ser estabelecida somente com o auxílio da experiência, 'dor = ativação da fibra-C' não é uma sentença necessária, assim como 'a = a' é uma sentença necessária. No caso das sensações e estados cerebrais seria plausível imaginarmos outras situações nas quais organismos que sentissem a sensação de tipo S não tivessem as fibras neurais de tipo F que nós temos. Nesses outros mundos possíveis, ter dor poderia ser outra coisa que ter a fibra-C ativada, diferentemente do que é o caso no nosso mundo possível.

Um problema característico a essa postura foi apontado por Kripke (1980). Os conceitos de analiticidade, aprioricidade e necessidade são estanques e não há entre eles relação de implicação mútua, diferentemente do que se poderia imaginar. Analiticidade é um conceito semântico, aprioricidade um conceito epistemológico e necessidade um conceito metafísico. Portanto, não podemos inferir a contingência das identidades, assim como aparentemente o fez Smart, do fato delas serem expressas por enunciados sintéticos a posteriori. Kripke, todavia, sugeriu que Smart estava certo em dizer que as identidades mente-corpo eram sintéticas e a posteriori. No entanto, se fossem verdadeiras, deveriam ser necessárias, isto é, verdadeiras em todos os mundos possíveis. Não examinaremos os pormenores desta crítica porque isso nos desviaria a uma discussão sobre externalismo semântico e a visão Kripkeana de mundos possíveis, digressão que exigiria muito mais espaço do que podemos aqui nos servir. No entanto, vale enfatizar que a contingência das identidades mente-corpo propostas por Smart é discutível.

A crítica que, no entanto, parece ter derrubado a teoria da identidade de tipo, da qual Smart foi um dos maiores defensores em pelo menos um momento de seu pensamento, adveio não de uma teoria causal da referência e do significado, mas sim, anteriormente às famosas palestras de

6 Faz-se importante enfatizar que quando dizemos que um tipo, ou propriedade, é idêntico a outro tipo, ou propriedade, estamos falando

imprecisamente; de fato, se são idênticos não são dois; é um só. Por isso, o que queremos dizer quando falamos que um tipo de estado mental é idêntico a um tipo de estado físico é simplesmente o seguinte: apesar de termos dois nomes, há somente um tipo, ou propriedade.

Kripke, de uma experiência de pensamento proposta por Putnam em vários de seus trabalhos (e.g., 1967). A idéia geral é simples: seria possível imaginarmos (em nossa situação ou mundo possível) que alguns organismos tivessem estados mentais sem terem as fibras neurais típicas daqueles estados mentais na raça humana. Moluscos, por exemplo, poderiam sentir dor sem terem a fibra específica da dor nos seres humanos. De modo mais geral, seria plausível que organismos de constituição física diferente da nossa tivessem sensações. O que, em princípio, impedir-nos-ia de considerar tipos de estados mentais realizáveis em múltiplos tipos de substratos físicos? Nada, aparentemente. Por conseguinte, ou a última possibilidade é apenas uma pseudo-possibilidade ou as identidades de Smart não se sustentam.

O Enfraquecimento das Identidades de Tipo

Smart certamente não foi o único a defender uma tal postura em relação ao debate mente-corpo: Place, Feigl, e Armstrong, entre outros, também defenderam hipóteses de identidade mente-corpo similares à de Smart. Apesar das diferenças de cada posição específica, o que se poderia considerar uma constante na proposta desses autores é a abrangência das identidades envolvidas: quando falam em identidades mente-corpo estão falando sobre identidades entre tipos mentais e tipos físicos, isto é, propriedades mentais são propriedades físicas. Portanto, se sentir coceira é ter a fibra X ativada, todo caso daquele tipo de sensação é um caso de ativação deste tipo de célula nervosa e vice-versa.

Assim como mencionado, Putnam, entre outros, avançou um ataque à teoria da identidade de tipo enfraquecendo-a de maneira implacável. Seu argumento tem como característica principal a abstração de propriedades mentais de propriedades físicas, de maneira que as primeiras sejam individuadas pelo seu papel funcional e não pela sua natureza física. Tomando-se o modelo de uma máquina de Turing como fonte de comparação, Putnam sugere que estados internos especificados em tabelas de máquina, determinados de acordo com suas funções correlativas, seriam análogos de estados mentais. O que eminentemente se ganha com essa analogia é precisamente a possibilidade de realização das mesmas propriedades mentais por propriedades físicas diferentes. Basta que dois sistemas realizem a mesma tabela de máquina para exemplificarem a mesma psicologia, independente de suas composições físicas (e.g., um poderia ser composto organicamente e o outro de cerâmica). Além disso, do ponto de vista metodológico, o trabalho do psicólogo, ou o do cientista da cognição, não se comprometeria a um nível físico específico, de modo que a autonomia da pesquisa psicológica pudesse ser novamente fortalecida. Isto é, o psicólogo preocupar-se-ia com *schemata* de leis psicológicas (viz., leis de tipo funcional) abstraídas de sistemas físicos específicos. É claro que em algum momento a ponte entre essas leis formais sobre a funcionalidade do mental e os substratos físicos dos sistemas cognitivos teria de ser estabelecida por meio de interpretações das descrições formais em termos de descrições físicas. Todavia, isso seria, presumivelmente, um

problema de ordem diferente do problema da pesquisa sobre a mente em geral.

Para que entendamos melhor a proposta geral do funcionalismo, vale discutirmos mais detidamente a noção de uma máquina de Turing. Uma tabela de máquina de Turing é uma descrição completa de um sistema possível, de sorte que indica as correlações entre *inputs*, *outputs* e estados internos desse sistema: para cada *input*, dado um estado interno, segue-se um estado interno e um *output*. Assim, poder-se-ia entender uma tabela de máquina como um quadro sinótico de funções matemáticas que atribuem uma ênupla ordenada (composta de *outputs* e outros estados internos) a cada argumento (entrada). De fato, cada estado interno de máquina seria ele próprio uma função matemática e a tabela seria a totalidade dessas funções e de suas possíveis atribuições. Um exemplo extremamente simplificado de uma tabela de máquina seria o seguinte (Block, 1980):

Input	S_1	S_2
50 centavos	Emite nenhum output Vai para S_2	Emite um guaraná Volta para S_1
1 Real	Emite um guaraná Permanece em S_1	Emite um guaraná e cinquenta centavos Volta para S_1

Esta tabela especifica as funções que regem as associações entre entradas, saídas e outros estados internos do sistema em questão, uma máquina de refrigerantes. S_1 , nesse caso, seria uma função de domínio {50 centavos, 1 real} e imagem {<nada, S_2 >, <guaraná, S_1 >}. Em outros termos, se colocamos uma moeda de 50 centavos na máquina [e ela está no estado S_1], ela passa de S_1 para S_2 e não emite nenhuma saída. Por outro lado, se colocamos uma moeda de um real na máquina [e ela está em S_1], ela emite um refrigerante e permanece em S_1 , e assim por diante em relação às outras duas possibilidades. Por conseguinte, a tabela nos mostra a relação dos possíveis comportamentos periféricos e mudanças internas do sistema, dadas certas entradas específicas e o estado interno do sistema no momento da entrada.

Notemos, porém, que na apresentação da tabela não jaz nenhuma especificação sobre a sua implementação física, de tal maneira que os processos computacionais por ela exibidos podem ser implementados em diversos substratos físicos. Similarmente, se entendemos o ser humano como uma máquina de Turing, ou um autômato probabilístico, podemos identificar seus estados psicológicos a estados internos de máquina e assim acomodar a intuição de que propriedades mentais são mentais não devido ao substrato ao qual se aplicam, mas devido à função que têm na economia computacional do sistema. Dores, coceiras, crenças, desejos, etc., seriam estados internos de máquina que correlacionariam certas entradas a certas saídas e outros estados internos. Por exemplo, o estado interno 'dor' correlacionaria (em uma versão vergonhosamente simples desse fenômeno) o dano no tecido do organismo em questão com a produção de comportamento de escape, gemidos e a sensação de angústia, se o organismo estiver alerta, etc. Portanto, o que caracterizaria o estado mental 'dor' seria o seu papel funcional face a

estímulos, respostas e outros estados, e não o fato de ser realizado fisicamente desta ou daquela maneira.

A diferença essencial entre o funcionalismo, exemplificado pelo funcionalismo de Máquina, e a teoria da identidade está na individuação das propriedades mentais. Enquanto que o último individua propriedades mentais a partir de propriedades físicas, pois propriedades mentais são propriedades físicas, o primeiro individua propriedades mentais identificando-as a propriedades funcionais—nesse sentido o funcionalismo também é uma espécie de teoria da identidade de tipo, pois identifica tipos mentais a tipos funcionais; porém, é importante que não confundamos isso com a teoria da identidade de tipo à maneira de Smart, que identifica tipos mentais não a tipos funcionais, mas a tipos físicos (para evitarmos confusões, usaremos a expressão ‘teoria da identidade de tipo’ no sentido de Smart). De volta à noção de propriedade funcional, em termos gerais, podemos dizer que uma propriedade funcional é uma característica que um dado sistema tem em virtude de se comportar de um certo modo dados de certos *inputs*. No caso do funcionalismo de máquina, propriedades funcionais são os estados internos do sistema, exibidos na tabela de máquina. No entanto, há outros tipos de funcionalismo que nos parecem oferecer propostas mais palpáveis em relação à noção de propriedade funcional. Uma das espécies de funcionalismo que se tornou bastante popular em filosofia da mente é o funcionalismo causal de tipo fisicalista. De acordo com essa postura, ter um estado mental é ter um estado físico tal que este (estado físico) seja causado e cause outros estados físicos. Por exemplo, ter dor é ter um tipo de fibra neural tal que é causada por outras ativações nervosas e causa outras, que em última análise causam o comportamento típico de dor. Nesse sentido de funcionalismo, ter mente é ter certas características físicas estruturadas causalmente de uma dada maneira.

No que concerne, todavia, ao problema da causação mental, o funcionalismo não parece particularmente atraente. Em primeiro lugar porque a metafísica funcionalista tende a ser um pouco imprecisa, ponto enfatizado por Kim em seus últimos trabalhos (e.g., 1998). Dizer que uma propriedade funcional é uma característica delimitada em virtude do papel causal de uma outra propriedade física em relação a uma rede de propriedades físicas é dizer que propriedades funcionais são propriedades de propriedades, isto é, propriedades de segunda-ordem tal que tê-las é ter certas propriedades de primeira-ordem (nesse caso físicas) em um certo arranjo causal (Kim, 1998). Por exemplo, sentir coceira (propriedade funcional de segunda-ordem) é, de acordo com o funcionalista fisicalista, ter um certo tipo de fibra nervosa ativada, de modo que essa ativação é causada, em última análise, por indutores de coceira e cause, em última análise, comportamento típico de coceira. Porém, o problema que nos preocupa fundamentalmente é o de como é possível uma propriedade mental, ela mesma, causar uma propriedade física. E isso o funcionalista não responde, pois a partir de sua concepção de propriedade mental como propriedade funcional individua-se estados mentais através do papel causal de outros estados, nesse caso, estados físicos. Voltando ao

exemplo da dor, ter dor, de acordo com essa concepção funcionalista-causal, é ter um estado físico tal que esse estado físico seja causado por, por exemplo, dano na pele, e cause comportamento típico de dor. Mas o que o filósofo da mente quer saber em relação ao problema da causação mental é precisamente qual o status causal de estados mentais (e.g., dores, coceiras, crenças, desejos, etc.), eles mesmos, diante do poder causal evidente dos estados físicos que os embasam. Quando digo que a minha vontade de beber água causou o meu movimento até a cozinha, quero dizer o quê? Que a minha crença de que há água potável na cozinha, em conjunto com o meu desejo de saciar a minha sede, causam o meu movimento corporal até a cozinha? Mas os estados cerebrais que embasam essas crenças e desejos certamente têm poder causal, o que acarreta dizer que tanto as crenças e desejos quanto os estados cerebrais têm poder causal. Mas como podem eles ter o mesmo poder causal e permanecer distintos? Em suma, a posição funcionalista em relação à natureza dos estados mentais deixa de responder perguntas essenciais em relação ao problema da causação mental.

O Monismo Anômalo

Em 1970, Donald Davidson apresentou uma solução alternativa para o problema da causação mental. Tal alternativa, o monismo anômalo, pretendia mostrar uma saída que ao mesmo tempo não se comprometesse com o fisicalismo de tipo nem com propostas abertamente dualistas (e.g., o cartesianismo, paralelismos de várias espécies, etc.). Desta maneira, evitar-se-iam tanto os problemas implicados pela aceitação de teses fisicalistas de tipo quanto as inconveniências acarretadas por teses dualistas.

Davidson abre o seu seminal trabalho (1970) expondo um problema aparentemente insolúvel: como seria possível conciliar o papel causal de estados mentais, dada a sua resistência à captura na rede causal física? A aparente contradição, diz-nos Davidson, emerge da conjunção de três teses: (i) o princípio da interação causal, (ii) o princípio do caráter nomológico da causalidade e (iii) o anomalismo do mental. O princípio da interação causal diz-nos que pelo menos alguns eventos, ou processos, mentais se relacionam causalmente com eventos físicos. O princípio da causalidade legiforme reza que onde há causalidade há lei, isto é, eventos relacionados como causa e efeito recaem sob leis estritas e determinísticas. No entanto, para a nossa surpresa, o princípio do anomalismo do mental propõe a idéia de que não há leis estritas e determinísticas a partir das quais possamos prever e explicar eventos mentais.⁷ Ora, isso parece inconsistente, pois se aceitamos que eventos mentais têm poder causal no mundo físico e aceitamos o caráter nomológico da

7 Essa noção de lei estrita e determinística é apresentada por McLaughlin (1989) da seguinte maneira: uma lei estrita e determinística é uma lei expressa no vocabulário de uma teoria (a) fechada e (b) abrangente, sendo que uma teoria é fechada se, e somente se, os seus eventos se relacionam causalmente somente entre si, e é compreensiva se, e somente se, interações causais entre seus eventos são subsumidas por leis dessa teoria.

causação, então não podemos aceitar o anomalismo do mental, pois esta última tese acarretaria a impossibilidade do poder causal de eventos mentais, certo? Errado. Segundo Davidson, é possível conciliarmos o poder causal de eventos mentais com seu anomalismo (i.e., o não-recaimento em leis estritas e determinísticas), e isso se faz a partir de uma teoria da identidade mente-corpo diferente de uma identidade de tipo.

A teoria da identidade mente-corpo defendida por Davidson tem como escopo somente eventos mentais e físicos. Este ponto é especialmente importante porque dependendo do que aceitamos sobre a natureza de eventos a teoria da identidade proposta por Davidson pode deixar de fazer sentido. Para que isso não aconteça, ou seja, para que o argumento de Davidson seja eficaz, ou pelo menos *prima facie* plausível, devemos entender eventos quaisquer como indivíduos, ou particulares, datados e não-repetíveis. Nesse sentido, uma erupção particular de um vulcão seria um evento, assim como o nascimento ou morte de uma pessoa. Dadas as condições de identidade dos eventos davidsonianos, expressões como 'A morte de Scott = A morte do autor de Waverley' referir-se-iam aos mesmos acontecimentos. Posto isso, Davidson defende a tese de que eventos físicos (e.g., a ativação de um conjunto de células nervosas em João, no instante de tempo *t*) são eventos mentais (e.g., a crença, em *t*, que João tem de que faz sol) isto é, são o mesmo acontecimento descrito de maneiras diferentes.

A tese da identidade proposta por Davidson é, com efeito, um *specimen* do que chamamos genericamente de teoria da identidade de ocorrência (*token identity theory*, em oposição à posição smartiana, *type identity theory*). Em sua forma geral, a teoria da identidade de ocorrência postula que ocorrências de estados mentais são ocorrências de estados físicos; e a proposta de Davidson, por meio da identidade de eventos mentais e físicos, acarreta exatamente isso: cada ocorrência de um estado mental é uma ocorrência de um estado físico. Note-se, porém, que consoante a posição davidsoniana as relações entre os tipos físicos e mentais são simplesmente ignoradas. E, na realidade, há de ser assim para que o argumento davidsoniano seja eficaz, pois é parte essencial de sua postura que não haja nenhuma relação mais forte com tipos, ou propriedades, mentais e físicas.

Para satisfatoriamente tratarmos o problema da causação mental, de acordo com Davidson, precisamos, além de uma teoria da identidade entre eventos mentais e físicos, da tese do anomalismo psicofísico. Enunciemos assim o anomalismo psicofísico: não há leis psicofísicas, no sentido de não haver correlações legiformes entre propriedades mentais e propriedades físicas. Substancialmente, o que essa tese diz é que não há correlações fortes ligando a psicologia à pesquisa do mundo físico, o que automaticamente salvaguarda a psicologia de qualquer redução interteórica ao físico. Mas, perguntar-se-á, qual a justificação para essa tese?

Uma resposta exaustiva para tal pergunta nos exigiria muito mais do que podemos aqui oferecer. No entanto, podemos, de maneira esquemática e geral, apresentar a intuição por trás do argumento davidsoniano. Começemos por

chamar atenção ao fato de que Davidson, quando discute mentalidade, se refere especificamente a estados intencionais. Nesse sentido, ter mente é ser descrito através de verbos como 'crer', 'desejar', 'temer', etc., de modo que as razões para os nossos atos são verbalizadas em forma de atitudes proposicionais. Portanto, ao dizermos que Davidson trata do problema mente-corpo, em especial do problema da causação mental, por meio do monismo anômalo, referimo-nos aos seus argumentos em relação ao poder causal de estados de tipo intencional. Já os estados qualitativos, como as sensações, seriam redutíveis a estados intencionais, o que faria de toda a mentalidade uma grande teia intencional.

Se por um lado a intencionalidade exaure o âmbito mental, por outro o mundo físico carece de toda e qualquer relação essencial com a intencionalidade. Deste modo, os domínios do mental e do físico seriam constitutivamente diferentes, sendo suas características essenciais absolutamente independentes. Mas o que, mais especificamente, diferencia o mental do físico? A resposta davidsoniana sugere que a atribuição de mentalidade a algo dá-se de acordo com certos princípios de racionalidade, coerência e consistência. Isso quer dizer que atribuir desejos, crenças e temores a outrem é parte de uma estratégia explicativa regida por princípios de maximização de racionalidade que têm como fim conferir inteligibilidade às suas ações. Na verdade, seria *prima facie* discutível atribuímos mentalidade a algo sem que essa atribuição se desse de acordo com princípios que salvaguardassem a consistência e a coerência dos estados mentais atribuídos, pois os estados intencionais têm função representativa, por meio de proposições; estas, por sua vez, têm relações de consistência e coerência determinadas entre si, holisticamente, devido ao seu conteúdo. Por exemplo, quando atribuímos a alguém a crença de que um café expresso custa 50 centavos, a regra da maximização da racionalidade nos impele a atribuímos também ao mesmo sujeito a crença de que o mesmo café, sob as mesmas condições, custe mais que 5 centavos ou mais do que 15 centavos, pois se assim não o fizermos correremos o risco de não conseguir interpretar as ações daquele sujeito. Se uma pessoa diz que gosta e não gosta do trabalho que o Presidente da República vem desenvolvendo, normalmente não tomamos isso como indicação de que a pessoa está emitindo uma clara contradição; ao contrário, a fim de que encontremos sentido nos atos e na mentalidade desta pessoa, interpretamo-la como dizendo que gosta de certos aspectos da administração do Presidente mas não de outros. Mesmo que a pessoa insista em proferir que gosta e não gosta de algo e ponto final, provavelmente procuraríamos uma maneira alternativa de entender o conectivo 'e' (ou o 'não'), para assim alcançarmos uma interpretação maximamente consistente da psicologia da pessoa em questão, a partir da escassa evidência a nós disponível. Posto isso, o que Davidson parece sublinhar é que o mundo mental tem características constitutivas que são definitivas do que é ser mental, e por isso mesmo inseparáveis da noção de mentalidade, ao passo que o mundo físico carece dessas características especiais e idiossincráticas do mental, exibindo outras características essenciais incompatíveis com as do mundo

mental. Como o diz explicitamente Davidson (1970), essas condições de atribuição de estados mentais não têm eco na teoria física. E aqui se mostra a impossibilidade de haver leis ligando esses dois domínios, pois se assim não fosse, isto é, se houvesse correlações legiformes ligando o mental ao físico, transmitir-se-iam as características essenciais de um domínio para o outro, quebrando-se assim a hegemonia dos princípios constitutivos de cada âmbito.

Analisemos o seguinte exemplo a fim de aclararmos o argumento de Davidson (Kim, 1985). Suponhamos que a atribuição de estados mentais quaisquer, m_1 e m_2 , a um ser qualquer seja plausível. Suponhamos ademais que há estados neurais desse mesmo ser, n_1 e n_2 , tal que sejam nomologicamente coextensivos a m_1 e m_2 , respectivamente. Nesse caso, a ocorrência de n_1 implica necessariamente a ocorrência de m_1 , e vice-versa, valendo o mesmo para m_2 e n_2 . Já os estados neurais n_1 e n_2 são descritos por termos teóricos de uma teoria física e têm, por isso mesmo, condições de atribuição próprias à teoria física. Conseqüentemente, a teoria física prevê o necessariamente suficiente, no âmbito físico, para a ocorrência de n_1 e n_2 , independente do que acontece no âmbito mental. Chamemos C_1 a condição de atribuição de n_1 e C_2 a condição de atribuição de n_2 . Pela teoria física temos:

- 1a) Necessariamente, se C_1 acontece, n_1 acontece;
- 1b) Necessariamente, se C_2 acontece, n_2 acontece.

Pelas leis psicofísicas acima estabelecidas temos:

- 2a) Necessariamente, m_1 acontece se, e somente se, n_1 acontece;
- 2b) Necessariamente, m_2 acontece se, e somente se, n_2 acontece.

Daí inferimos:

- 3) Necessariamente, se C_1 acontece, m_1 acontece;
- 4) Necessariamente, se C_2 acontece, m_2 acontece.

Ora, o problema com essas inferências é que as atribuições dos estados mentais m_1 e m_2 tornam-se abertas à jurisdição do âmbito físico! Isto é, para que atribuamos estados mentais a outrem, basta que tenhamos a teoria física e as correlações psicofísicas legiformes relevantes (Kim, 1985). E isso é atribuir mentalidade independente das características constitutivas do mental. Seria como se pudéssemos, a partir da neurofisiologia, mapear toda a intencionalidade do sujeito em questão. No entanto, isso é precisamente o que Davidson não pode aceitar, pois a psicologia, de acordo com o seu entendimento de psicologia, não é, e não pode ser, determinada exhaustivamente pela neurofisiologia. Logo, se aceitamos as infames leis psicofísicas, perdemos em nossa caracterização a essência do mental e, por isso, perdemos o mental, pelo menos no sentido para Davidson relevante de mental (Kim, 1985). Isso, contudo, não quer dizer que não possa haver generalizações verdadeiras *ceteris paribus* correlacionando os âmbitos mental e físico; essas últimas generalizações não dispõem da força modal que as leis dispõem, tornando assim as correlações psicofísicas inócuas no que tange aos princípios constitutivos de cada âmbito. E é exatamente por isso que o argumento de Davidson vale—se é que podemos dizer que é válido—somente para leis

psicofísicas, e não para generalizações *de facto* entre o mental e o físico.

O que fundamenta, então, a tese do anomalismo psicofísico (viz., não há leis psicofísicas) é precisamente a diferença de compromisso teórico entre o mental e o físico e a autonomia teórica do mental em relação ao físico e vice-versa. Como vimos, as dificuldades surgem se os princípios constitutivos do mental repartem sua jurisdição com os princípios do domínio do físico, tornando as atribuições de estados intencionais passíveis de explicação suficiente por princípios físicos, fato este que descaracterizaria o mental como mental. Assim, se fossem possíveis as leis psicofísicas, falar do mental seria o mesmo que mudar de assunto, como nos diz Davidson (1970).

Suponhamos a validade do argumento para o anomalismo psicofísico. Mesmo assim, indagar-se-á, como é possível chegarmos à tese do anomalismo do mental (viz., propriedades, ou descrições, mentais não podem ocorrer em leis estritas e determinísticas)? Lembremos que as duas teses são diferentes: o anomalismo psicofísico proíbe correlações legiformes entre propriedades, ou predicados, mentais e físicos; já o anomalismo do mental é mais geral, pois sustenta que propriedades, ou predicados, mentais não podem ocorrer em leis estritas quaisquer. A pergunta, então, se põe: como é possível inferir do anomalismo psicofísico o anomalismo do mental? Kim propõe (1985) que aceitemos a tese adicional do anomalismo psicológico. Tal tese indicaria que nem mesmo no âmbito restrito da intencionalidade poderia haver correlações legiformes. Por conseguinte, leis psicológicas seriam impossíveis e, com isso, a possibilidade de cientificação da psicologia estaria descartada, o que, enfatizamos, é consistente com o projeto davidsoniano (e.g., Davidson, 1974). Todavia—insistir-se-á—como pode o âmbito psicológico ser anômalo se formulam-se leis psicológicas, leis estas que caracterizam exatamente os compromissos constitutivos do mental? Em outros termos, se precisamos de uma *teoria* para a atribuição de estados intencionais a outrem, como é possível termos isso sem termos ao mesmo tempo leis psicológicas? A resposta de Davidson diz que as leis da psicologia são normativas, e não explicativas e preditivas como as leis do domínio físico (Kim, 1985). Destarte, tais leis do domínio mental não teriam como objetivo senão a interpretação das ações humanas, projeto este que se diferencia claramente dos propósitos das ciências naturais enquanto empreendimento que visa a explicação baseada na causalidade. As leis normativas da psicologia teriam como meta o exame das razões para a ação humana, procedimento que se assemelha muito mais à hermenêutica do que à previsão e explicação científica propriamente ditas (Kim, 1985). Assim, o princípio do anomalismo psicológico seria válido, pois consistente com essas “leis” normativas características da explicação psicológica. Já a tese geral do anomalismo do mental seria exatamente a conjunção do anomalismo psicofísico com o anomalismo psicológico.

A partir disso, o que podemos dizer do poder causal dos eventos mentais, isto é, como podem estes causarem algo se são anômalos, e se toda relação de causalidade é subsumida

por uma lei estrita? A resposta é clara: eventos mentais têm poder causal *qua* físicos e não *qua* mentais. Como todo evento mental tem uma descrição física (ou partilha de uma propriedade mental, no linguajar da metafísica), então o seu poder causal deve-se ao fato do evento recair em descrições (ou propriedades) físicas. Eventos mentais, enquanto descritos fisicamente, são passíveis de subsunção por leis estritas e determinísticas. Assim, salvamos a causalidade de eventos mentais e, ao mesmo tempo, a autonomia do mental, resolvendo o problema da aparente contradição entre o anomalismo do mental e seu poder causal (i.e., de seus eventos). Temos um monismo de eventos (todos os eventos são físicos) e um anomalismo dos tipos, ou propriedades, mentais.

Se a resposta de Davidson, por um lado, responde ao problema do poder causal das ocorrências mentais, por outro deixa de tratar a possibilidade de relação entre os tipos mentais e físicos, subscrevendo implicitamente a um dualismo de tipos. Assim, se pretendemos alcançar uma posição satisfatória em relação ao problema de como estados mentais, enquanto mentais, causam estados físicos, o monismo anômalo não nos parece a melhor saída.

A Superveniência e o Argumento de Kim

A longo desse texto investigamos, em linhas gerais, o que se entende atualmente pelo problema mente-corpo em filosofia da mente, em particular no que diz respeito ao problema da causalidade mental. Assim como vimos, a identidade mente-corpo exemplificada na proposta de Smart foi bastante enfraquecida pela possibilidade de múltiplas realizações de estados mentais; o argumento de Putnam mostrou-se dramaticamente pungente em relação à possibilidade de uma redução mente-corpo a partir da identidade de tipos mentais e físicos. Já o funcionalismo, de maneira geral, admitiu a múltipla realização do mental pelo físico, talhando os tipos mentais a partir de seus papéis causais-funcionais na economia dos sistemas cognitivos. Apesar da plausibilidade inicial do funcionalismo, um problema crucial que emerge desta posição sobre a individuação de tipos, ou propriedades, mentais codifica-se na seguinte pergunta: se quem tem poder causal em um sistema cognitivo é, em última análise, um estado físico, qual o papel do estado mental em uma rede causal física?

O monismo anômalo, por sua vez, lida exclusivamente com o problema da causalidade de ocorrências mentais. Davidson faz de todos os eventos mentais eventos físicos, e, assim, explica o poder causal de ocorrências mentais por estas recaírem em propriedades físicas, que por sua vez podem ocorrer em leis estritas e determinísticas. Desta maneira, eventos mentais têm poder causal *qua* físicos e não *qua* mentais. Méritos ou deméritos das estratégias argumentativas de Davidson à parte, mesmo que a resposta davidsoniana seja boa, no sentido de não haver nenhuma inconsistência embutida no seu próprio argumento, ela carece, assim como insiste Kim (1998), de uma apresentação de uma tese positiva sobre a relação entre tipos mentais e físicos. Davidson

aparta os domínios físico e mental de tal modo que a mentalidade tomada enquanto tal mostra-se absolutamente supérflua na rede causal física (Kim, 1989). Com efeito, as seguintes considerações confirmam essas intuições. Poder-se-iam redistribuir todas as propriedades mentais de maneira diferente do que agora é o caso e ainda assim poderíamos ter um mundo físico causalmente idêntico ao nosso. Pior ainda, poder-se-ia eliminar completamente o âmbito do mental sem que a rede causal física sofresse qualquer mudança (Kim, 1989). Portanto, até onde entendemos os argumentos de Davidson, a sua resposta parece insatisfatória para dar conta da existência da mentalidade.

Uma outra tentativa de resolução do problema se deu a partir da idéia de superveniência. A superveniência, em geral, é uma relação de covariância entre (conjuntos de) propriedades no sentido em que se houver indiscernibilidade no nível subveniente, há indiscernibilidade no nível superveniente (cf., Kim, 1990). Assim, dizer que a mente supervém ao físico é dizer que dados x e y (em quaisquer mundos possíveis), se x e y têm as mesmas propriedades físicas, necessariamente x e y têm as mesmas propriedades mentais.⁸ Além de aclarar a relação entre propriedades mentais e físicas, oferecendo-nos um modelo que capturasse as regularidades entre os âmbitos físico e mental, a relação de superveniência também aparentava oferecer uma resposta para o problema da causalidade mental: propriedades mentais causariam propriedades físicas supervenientemente, isto é, em virtude de suas bases físicas causarem as mesmas propriedades físicas. Kim chegou a defender uma posição desse tipo em alguns de seus trabalhos (e.g., 1984). No entanto, tal posição se mostrou insustentável, pois supunha a duplicação de cadeias causais desembocando no mesmo evento físico. O seguinte argumento apresentado por Kim (1998) elucida muito bem o núcleo do problema:

$M \text{ — } M^*$

$P \text{ — } P^*$

1. Suponhamos que uma propriedade mental, M , cause uma outra propriedade mental, M^* . Por exemplo, a crença de que está frio causa o desejo de usar um agasalho. De acordo com a noção de superveniência apresentada, propriedades mentais supervém a propriedades físicas no sentido em que se algo tem uma propriedade mental, M ,

⁸ Uma outra maneira de se apresentar a noção de superveniência seria a seguinte: Um conjunto, M , de propriedades supervém a um conjunto, F , se, e somente se, necessariamente para qualquer x que tenha propriedade m em M existe uma propriedade f em F , tal que x tem f , e, necessariamente, qualquer coisa que tenha f tem m . (Kim, 1997). Com qualificações, essas duas apresentações da noção de superveniência são equivalentes (Kim, 1997, cf., Savellos & Yalçın, 1995). A primeira versão é chamada 'versão da indiscernibilidade' e a segunda 'a versão do operador modal'. Há também diferenças de força modal entre essas definições. As duas que apresentamos aqui são versões fortes, pois exigem a covariância de propriedades em todos os mundos possíveis.

no tempo t , há uma base física, P , tal que esse algo tem P , e necessariamente qualquer coisa que tem P no tempo t , tem M em t (Kim, 1998). Assim,

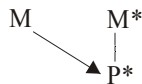
2. M^* tem uma base física P^*

Mas o que explica a ocorrência de M^* ? Duas são as possíveis respostas:

3. (3.1.) Por hipótese, porque M causou M^* ; (3.2.) pela tese da superveniência mente-corpo, porque P^* , a base física de superveniência de M^* , ocorre nessa ocasião.

Vale lembrarmos que segundo a tese da superveniência, a base, P^* , é suficiente para a ocorrência de M^* . Por isso, basta que P^* ocorra para M^* ocorrer, independente do que possa acontecer além disso. Desta maneira, a ocorrência de M parece supérflua, pois se P^* ocorre, M^* ocorre, chova ou faça sol. Logo, o poder causal de M em relação a M^* encontra-se ameaçado, sendo a sua única possível reconciliação (sem admitirmos de imediato a exclusão de M como fator causal):

4. M causa a ocorrência de P^* , que por sua vez é suficiente para a ocorrência de M^* .



Com efeito, essa proposta nos mostra algo interessante no esquema da superveniência mente-corpo: se quisermos afetar as propriedades supervenientes, temos que passar primeiramente pelas propriedades subvenientes, pois estas são necessariamente suficientes para as ocorrências das propriedades supervenientes.⁹ No entanto, o problema com o passo (4) do argumento é que temos um exemplo de causação descendente, isto é, um exemplo no qual uma propriedade mental causa uma propriedade física. Mas isso é exatamente o que queremos evitar! Em que sentido podemos dizer que uma propriedade mental *causa* uma propriedade física? Ao aceitarmos a causação descendente, não estaríamos diante de uma espécie de interacionismo cartesiano, só que desta vez no domínio das propriedades? Como podemos dar conta da idéia de que algo não físico causa algo físico? Além disso, M , consoante a superveniência mente-corpo, pressupõe uma propriedade base, a saber, P . Assim,

5. M ocorre na ocasião da ocorrência de P , que é suficiente para a ocorrência de M .

Vejamus o seguinte diagrama (cf., Kim, 1996):



⁹ Lembramos também que o modelo de superveniência é consistente com a múltipla realização. As propriedades base, nesse sentido, podem mudar de uma espécie para outra, ou até mesmo de um indivíduo da mesma espécie para outro. O que parece fundamental, porém, é que para que uma propriedade mental qualquer ocorra, é preciso que haja uma propriedade física que a ocasione, de tal modo que a propriedade física seja suficiente para a ocorrência da propriedade mental.

Sejam M a crença de que está frio e M^* o desejo de se proteger do frio. P e P^* são ativações neurais ocasionando, respectivamente, M e M^* . A pergunta a se fazer é a seguinte: em que sentido a crença de que está frio *causa* o desejo de proteção?

Recapitulando, se M ocorre, há uma base P que a ocasiona; se M^* ocorre, há uma base P^* que a ocasiona. Como P e P^* são propriedades físicas, então não há nenhum problema especial na admissão de uma relação de causalidade entre elas. No entanto, nossa intuição inicial quanto ao poder causal de M em relação a M^* mostra-se em sérias dificuldades em dois momentos: (I) M parece supérflua, dada a suficiência de P^* para a ocorrência de M^* ; (II) por isso, se M tem poder causal sobre M^* , tal se deve, quiçá, ao poder causal que M tem sobre P^* , que por sua vez é suficiente para M^* . Só que, neste caso, há causação descendente (viz., de M a P^*).¹⁰ Porém, mesmo que aceitássemos que M causa P^* , aceitando assim a causação descendente, não se poderia ignorar que M , ela mesma, tem uma base subveniente P . Mas se P é suficiente para M e M é suficiente para P^* , então P é suficiente para P^* , por transitividade. Se entendemos a causação como fundada na noção de suficiência nomológica, P qualifica como causa para P^* (Kim, 1998). Alternativamente, se preferimos entender causação em termos de contrafactuais também temos ótimas razões para acreditar que P causa P^* , pois se P não ocorre, M não ocorre, e se M não ocorre, P^* não ocorre (Kim, 1998). (Isso não quer dizer, entretanto, que defendamos a existência de uma relação de causalidade entre P e M , pois isso é exatamente o que negamos; porém, tudo o que queremos mostrar é que *se* entendemos M como causa de P^* , então teremos que aceitar que P também causa P^*). A moral da história é que a causação descendente gera a superabundância causal; cada evento cerebral tem tanto uma causa física como uma causa mental! Se nossa argumentação faz sentido, uma saída aparentemente mais atraente parece ser aquela de acordo com a qual M não tenha poder causal frente ao poder causal de sua base subveniente, P .

6. As regularidades M -a- M^* e M -a- P^* são derivativas (i.e., causação superveniente, ou pseudo-causação), surgindo de um processo causal genuíno entre P e P^* .

Porém, se aceitamos que a mentalidade seja de fato epifenomênica ao mundo físico, isto é, sem poder causal algum diante do físico, também não obtemos uma resposta cabal para o problema da causação mental. Com efeito, o problema em assumir-se tal postura é que fracassamos em estabelecer a realidade da mente em um mundo físico. Se concordamos com Samuel Alexander sobre o que é ser real (viz., ser real é ter poder causal), então não faz sentido falarmos que a mentalidade é real e epifenomênica.

¹⁰ Note-se que em relação à causação descendente, como qualquer posição em filosofia, precisamos de *argumentos* em seu favor. Simplesmente *dizer* que há causação da mente para o corpo não é o bastante! O que Descartes fez foi dizer que havia interação mente-corpo e o que os seus críticos rapidamente notaram é que faltava a especificação de *como* isso poderia se dar, dadas as premissas do argumento cartesiano.

O fiscalista não-reduutivo encontra-se, por conseguinte, em uma situação dramática: ou perde a realidade do mental ou abdica de sua posição não-reducionista. As duas teses tomadas em conjunto mostram-se dificilmente conciliáveis, situação esta que nos chama a atenção para a decisão de qual tese responsabilmente descartar. Kim, nos últimos anos, enfatiza que dentro dos limites do fiscalismo, a mais aceitável proposta é o reducionismo, de acordo com o qual propriedades mentais seriam propriedades físicas (em espécies ou estruturas específicas). Manter-se-ia a linguagem do mental, excluindo-se a ontologia permissiva característica do dualismo de propriedades. Tal projeto, ademais, estaria intimamente ligado a uma crítica dirigida ao funcionalismo de teor fiscalista. Na realidade, o núcleo da nova posição de Kim (1998) jaz na superposição do problema explicitado pelo argumento da superveniência na postura funcionalista fiscalista.

Seja lá como for, a relação de superveniência, antes considerada uma solução para o problema da causalização mental, simplesmente nos apresenta o problema (Kim, 1998). Como mencionado, temos duas opções: o epifenomenismo ou a redução. Se aceitamos a primeira, propriedades mentais deixam de ser propriedades causais, e por isso mesmo deixam de ser nômicas. Por conseguinte, a possibilidade de uma ciência do mental seria eliminada em princípio, pois se propriedades mentais não podem figurar em leis, como seria possível fundar-se uma pesquisa científica do mental? (Kim, 1998) Ademais, se propriedades mentais são meros epifenômenos, o que dizer da possibilidade do livre arbítrio? Dado que nossas crenças e desejos não têm influência nenhuma em nosso comportamento, em que sentido podemos dizer que somos livres? (Kim, 1998). Por outro lado, temos a redução que também envolve várias dificuldades, entre elas, a especificação de um programa teórico que explicitamente como estados intencionais e sensoriais possam ser entendidos como estados físicos.

Portanto, a situação na metafísica da mente contemporânea não nos parece das melhores. Uma das coisas que, no entanto, pretendemos mostrar com este trabalho é que diferentemente do que foi por muitos considerado, o problema mente-corpo está vivo. Se pretendemos entendê-lo, e quem sabe resolvê-lo, um ótimo ponto de partida parece ser o núcleo do argumento da superveniência: se para cada estado mental há uma base física que o ocasiona, e as bases físicas estão todas ligadas causalmente, como é possível que os estados mentais, enquanto mentais, tenham influência na rede causal física? Com efeito, essa pergunta se parece muito com a pergunta feita a Descartes sobre o mesmo problema! Mesmo sem o dualismo de substâncias, padecemos de um problema similar. E a resolução de tal problema, assim como nos indica Kim, parece de qualquer maneira dolorosa. Se há algo, em geral, a ser dito pelo reducionismo é o seguinte: se o aceitamos pelo menos não desistimos da possibilidade da causalização mental, assim como o faria um epifenomenista. Agora, uma defesa ampla e sistemática de uma posição reducionista estaria, sem dúvida alguma, além dos limites deste trabalho.

Referências

- Block, N. (1980). Introduction: What is Functionalism?. Em Ned Block (Org.), *Readings in the Philosophy of Psychology: Vol. I*. Cambridge (MA): Harvard University Press.
- Davidson, D. (1970). Mental Events. Em Lawrence Foster & J.W. Swanson (Orgs.), *Experience and Theory*. Amherst: The University of Massachusetts Press.
- Davidson, D. (1974). Psychology as Philosophy: Comments and Replies. Em S.C. Brown (Org.), *Philosophy of Psychology*. London: The Macmillan Press.
- Kim, J. (1984). Epiphenomenal and Supervenient Causation. *Midwest Studies in Philosophy* 9, 257-270.
- Kim, J. (1985). Psychophysical Laws. Em Ernest LePore & Brian McLaughlin (Orgs.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. (pp. 369-386). Oxford: Basil Blackwell.
- Kim, J. (1989). The Myth of Nonreductive Materialism. *Proceedings and Addresses of the American Philosophical Association* 63, 31-47.
- Kim, J. (1990). Supervenience as a Philosophical Concept. *Metaphilosophy* 21, 1-27.
- Kim, J. (1992). Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research* 52, 1-26.
- Kim, J. (1993). The Nonreductivist's Troubles with Mental Causation. Em John Heil & Alfred Mele (Orgs.), *Mental Causation*. (pp. 189-210). Oxford: Oxford University Press.
- Kim, J. (1996). *Philosophy of Mind*. Boulder: Westview Press.
- Kim, J. (1997). The Mind-Body Problem: Taking Stock after Forty Years. *Philosophical Perspectives* 11, 185-207.
- Kim, J. (1998). *Mind in a Physical World*. Cambridge (MA): MIT Press.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge (MA): Harvard University Press.
- McLaughlin, B. (1989). Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical. *Philosophical Perspectives* 3, 109-135.
- Putnam, H. (1967). The Nature of Mental States. Em W.H. Capitan & D.D. Merrill (Orgs.), *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press.
- Savellos, E. & Yalçin, U. (1995). Introduction. Em Ellias E. Savellos & Ümit D. Yalçin (Orgs.), *Supervenience: New Essays*. Cambridge: Cambridge University Press.
- Smart, J.J.C. (1959). Sensations and Brain Processes. *Philosophical Review* 68, 141-156.

Recebido em 25.02.2002

Aceito em 06.03.2002 ■