

CDD: 128.2

## FUNCIONALISMO E CAUSAÇÃO MENTAL

PAULO ABRANTES

*Departamento de Filosofia,  
Universidade de Brasília,  
70910-900 Brasília, DF  
BRASIL*

*abrantes@unb.br*

FELIPE AMARAL

*843 60th Street, Apt C2,  
Brooklyn, NEW YORK 11220  
USA*

*famaral@gc.cuny.edu*

*Resumo:* O que colocou o funcionalismo no centro do debate em torno do problema mente-corpo nos últimos trinta anos parece ter sido a sua capacidade de conciliar intuições fisicalistas com uma espécie de não-reducionismo: se por um lado postula-se a existência de entes físicos somente, distribuídos em uma ontologia estratificada, por outro não se falha em explicitar uma distinção real entre as propriedades de entes capacitados a sentir e representar. A superveniência mente-corpo aparentava esclarecer essas intuições dos fisicalistas não-redutivos. Vários dos trabalhos de Kim em torno dessa relação, em especial aqueles publicados antes do fim da década de oitenta, sugeriram a superveniência como uma possível, e promissora, elucidação do estado de coisas em tor no do problema da relação entre a mentalidade e o físico. Nos últimos anos, contudo, Kim apresentou um dilema, expressando com bastante pungência que a causalção mental revela-se ininteligível, valendo ou não valendo a superveniência mente-corpo. Por conseguinte, a superveniência deixa de compor uma possível solução para o problema, passando a constituí-lo. Como tal conceito de superveniência mente-corpo estende-se naturalmente a uma região de interseção entre todas as posturas não-reducionistas minimamente fisicalistas, as preocupações de Kim, se de fato justificadas, sugerem algo de bastante grave com o fisicalismo não-reducionista. Defendermos neste trabalho que, apesar do problema da exclusão causal ser genuíno, ainda é cedo para concluirmos a falsidade do fisicalismo não-redutivo.

*Palavras-chave:* Kim; fisicalismo; causalção mental; funcionalismo; mente-corpo.

A crítica funcionalista à teoria da identidade de tipo [*type*] parece ter sido a razão principal para o declínio desta última. Para os primeiros funcionalistas, como Putnam (e.g., (1965), (1967)) e Fodor (1965), a possibilidade da múltipla realização do mental se mostrava como uma intuição crucial, a ser preservada por qualquer presumível solução para o problema mente-corpo. A idéia de que mentes seriam análogas a máquinas computacionais implementadas em diferentes tipos de substrato mostrava-se bastante atraente, não somente porque consistente com, e sugerida pela pesquisa empírica sobre a cognição, mas também por fortalecer a autonomia da pesquisa do mental face à pesquisa do físico (Kim (1998)). Para o teórico da identidade, como se sabe, cada tipo mental é individuado por um tipo físico, engessando-se assim as relações mente-corpo e impossibilitando a múltipla realização, situação que desvela a inépcia desta postura para lidar com o problema mente-corpo.

Para responder por essa intuição da múltipla realização do mental, vários funcionalistas caracterizaram os estados mentais de um sistema cognitivo em termos da função causal desses estados numa rede causal envolvendo outros estados mentais, *inputs* e *outputs* do sistema.

Embora não se possa afirmar um compromisso necessário entre funcionalismo e fisicalismo, muitos funcionalistas são fisicalistas, em um sentido menos amplo de ‘fisicalismo’ do que o expresso por ‘fisicalismo de tipo’. Assim, estados mentais para estes funcionalistas podem ser realizados por estados físicos somente, apesar dessa dependência não acarretar a identificação do nível funcional ao nível físico.

O que parece ter posto o funcionalismo no centro do debate em torno do problema mente-corpo nos últimos trinta anos foi o fato dele ter se apresentado como uma “terceira via” entre dualismos de substância e fisicalismos de tipo. Destarte, o funcionalismo pareceu conformar-se a algumas das intuições fundamentais de doutrinas fisicalistas sem, com isso, abrir mão do não-reducionismo do mental ao físico. Se por um lado o funcionalista fisicalista postula a existência de entes físicos somente, distribuídos em uma ontologia estratificada, por outro, não falha em

explicitar a realidade e autonomia de certas características de entes físicos capacitados a sentir e representar.

A superveniência mente-corpo foi inicialmente vista por muitos como uma relação que esclarecia, justamente, essas intuições dos fisicalistas não-reducionistas. Vários dos trabalhos de Kim em torno dessa relação, em especial aqueles publicados antes do fim da década de oitenta, sugeriram a superveniência como uma possível, e promissora, elucidação do estado de coisas em torno do problema da relação entre a mentalidade e o físico. Contrariando essas expectativas, contudo, ao início da década de noventa Kim passou a sugerir com progressiva ênfase a insuficiência do conceito de superveniência para especificar inequivocamente uma relação de dependência entre o físico e o mental; a superveniência servir-nos-ia senão como mera elucidação de *padrões de covariância* exibidos entre o mental e o físico (e.g., Kim (1990)).

A relação de superveniência mostrou-se, além disso, incapaz de sustentar uma relação de causação genuína no nível do mental. Kim, nos últimos anos, apresentou um “dilema” diante do qual o fisicalista não-reducionista parece se deparar: para o não-reducionista, a causação mental revela-se ininteligível, valendo ou não valendo a superveniência mente-corpo. Como tal conceito de superveniência mente-corpo parece se estender naturalmente a uma região de intersecção entre todas as posturas não-reducionistas minimamente fisicalistas, as preocupações de Kim, se de fato justificadas, sugerem algo de bastante grave, em geral, com o não-reducionismo fisicalista em filosofia da mente.

O nosso propósito neste trabalho é modesto. Apresentaremos, em primeiro lugar, o dilema que parece assolar o fisicalismo não-reducionista. Em segundo lugar, sobreporemos o dilema sobre a tese do realizacionismo físico. Em terceiro lugar, compararemos a abordagem de Kim, em relação à natureza das propriedades mentais, com análises conceituais que parecem sugerir resultados análogos no que diz respeito à individuação de propriedades mentais em contextos restritos. Em quarto lugar, apresentaremos a tentativa de David (1997) em tratar o problema

da exclusão causal e sugeriremos que esta não se mostra muito promissora. Por fim, concluiremos que o ônus argumentativo está, neste momento, sobre os ombros do não-reducionista e, em particular, do funcionalista de persuasão fiscalista. Enfatizamos, porém, que isso não é o mesmo que dizer que o argumento da exclusão causal estabelece a falsidade do fiscalismo não-reducionista e, portanto, a falsidade do funcionalismo fiscalista; simplesmente, defenderemos que nenhum argumento convincente parece ter sido oferecido contra o problema da exclusão causal, nem mesmo através do funcionalismo fiscalista. Assim, aquela que parecia uma solução mostra-se por demais problemática sob a luz de uma análise metafísica mais detida.

## 1. O DILEMA DE KIM

Mencionamos dois pontos preliminares à apresentação do dilema:

(1) Há duas espécies de argumento sugeridas pelo dilema de Kim: argumentos contra o poder causal autônomo de *propriedades* mentais, entendidas como características mentais autônomas em relação ao físico, e argumentos contra o poder causal autônomo de eventos mentais, entendidos como exemplificações de propriedades mentais. Em geral, consideraremos a primeira espécie de argumento nuclear, apesar do dilema ser apresentado aqui em termos de eventos mentais e físicos, com o intuito de aumentar a sua inteligibilidade.

(2) Um segundo ponto preliminar diz respeito à existência de universais. Não endossaremos aqui teses realistas no que diz respeito aos universais, pois o dilema e, por conseguinte, a exclusão causal não perdem força caso endosseemos uma ontologia mais enxuta, livre de universais. O que parece crucial na apresentação do dilema é a diferença causal que características mentais fazem em um mundo físico, sejam essas características universais ou não-universais. Tais características, por exemplo, poderiam ser entendidas como *tropes*, ou propriedades particularizadas, interpretação que não diminui a importância da pergunta,

“como podem eventos mentais causarem eventos mentais ou físicos *em virtude de suas características mentais?*”. Assim, a seguinte tese é tudo o que está subentendido na nossa discussão do dilema de Kim: o poder causal que eventos têm depende de suas características, ou propriedades. Enfatizamos, todavia, que, apesar dessa tese ser consistente com teses realistas em relação aos universais, aquela não acarreta esta.

O dilema que explicita as dificuldades da superveniência psicofísica diante do problema da causação mental resulta de um argumento que possui os seguintes passos:

(1) A superveniência mente-corpo vale ou falha.

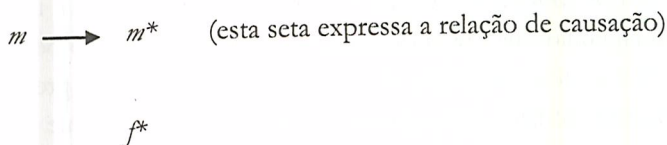
(2) Se a superveniência mente-corpo falha, não há como entender a possibilidade da causação mental. Se a superveniência mente-corpo vale, temos o que se segue.

(3) Suponhamos que um evento mental,  $m$ , cause um outro evento mental,  $m^*$ .

Por exemplo, suponhamos que a sensação de sede de um certo sujeito em um certo tempo, etc., seja a causa do desejo deste mesmo de saciar a sua sede em um certo tempo, etc. Nesse sentido, a ocorrência daquela sensação parece ser condição ao menos suficiente para a ocorrência desse desejo. Em outras palavras, o evento  $m$  parece causar o evento  $m^*$ .

De acordo com a superveniência mente-corpo, propriedades mentais supervêm a propriedades físicas no sentido em que se um objeto tem qualquer propriedade mental  $M$ , no tempo  $t$ , há uma propriedade física  $F$ , tal que esse objeto tem  $F$ , e, necessariamente, qualquer coisa que tenha  $F$  no tempo  $t$ , tem  $M$  em  $t$ . Assim, um evento de tipo  $M^*$ , a saber  $m^*$ , em  $t$  – o desejo de saciar a sede no nosso exemplo – pressupõe a ocorrência de uma propriedade física em  $t$  que a embase, um evento de tipo  $F^*$ , isto é,  $f^*$ . Logo,

(4)  $m^*$  tem uma base física  $f^*$ .



Cabe, no entanto, a seguinte questão: o que *explica*  $m^*$ ? Duas parecem ser as possíveis respostas.

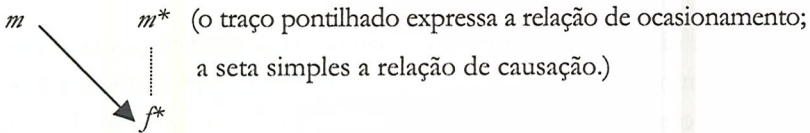
(5)  $m^*$  ocorre nesta ocasião: (5.1.) por hipótese, porque  $m$  causou  $m^*$ ; (5.2.) alternativamente, consoante a superveniência psicofísica, porque o evento  $f^*$  ocorre nessa mesma ocasião.

Parece importante lembrar que segundo a superveniência psicofísica, a base física,  $f^*$ , é *suficiente* para a ocorrência de  $m^*$ , e, além disso, *uma* base física é necessária para  $m^*$ . Por isso, basta  $f^*$  ocorrer para  $m^*$  ocorrer, e toda vez que  $m^*$  ocorre, uma base física ocorre, independente do que possa acontecer além disso. Destarte, a ocorrência de  $m$  parece supérflua, pois não é nem necessária, nem suficiente para  $m^*$ . Portanto, o poder causal de  $m$  em relação a  $m^*$  encontra-se ameaçado, sendo a seguinte a única saliente possibilidade de manutenção de  $m$  como fator causal:

(6)  $m$  causa  $f^*$ , que por sua vez é suficiente para  $m^*$ .

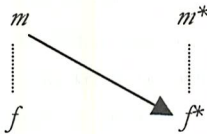
Com efeito, como Kim enfatiza em diversas ocasiões (e.g., (1998), p. 42), essa proposta nos sugere algo importante no esquema da superveniência mente-corpo: se quisermos afetar de alguma maneira eventos mentais, temos que passar primeiramente pelos eventos físicos que os embasam. Tal princípio salvaguarda um dos desideratos do programa fiscalista em filosofia da mente, que diz: mesmo que fenômenos mentais sejam diferentes de fenômenos físicos, os primeiros são necessariamente embasados pelos últimos, no sentido em que, se há alguma mudança no nível do mental, há mudança no nível que o embasa, o físico. Assim, voltando ao nosso modelo, se pretendemos salvar a

potência causal de  $m$ , face à suficiência de  $f^*$  para  $m^*$  (e a necessidade de eventos físicos para  $m^*$ ), parece-nos plausível consentir que  $m$  influencie  $m^*$  indiretamente, causando  $f^*$ , que por sua vez ocasiona  $m^*$ . O seguinte diagrama exhibe essa possível situação:

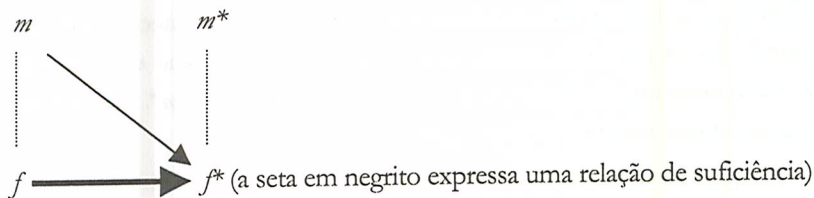


Não obstante a sua plausibilidade inicial, tal saída traz consigo ao menos uma dificuldade: a causação descendente. Em que sentido podemos dizer que um evento mental causa um evento físico? Não estaríamos aqui diante de uma espécie de interacionismo cartesiano? Além disso, como a superveniência psicofísica nos ensina, o evento  $m$  pressupõe ele mesmo uma base física. Assim,

(7) O evento  $m$  ocorre na ocasião do evento  $f$ , que é *suficiente* para  $m$ .



Por transitividade, se  $f$  é suficiente para  $m$ , e  $m$  é suficiente para  $f^*$  (por hipótese não-reducionista), então  $f$  é suficiente para  $f^*$ , o que gera a seguinte situação:



Esse diagrama parece capturar algumas conseqüências da asserção de que eventos mentais têm poderes causais independentes dos eventos físicos por eles ocasionados. Recapitulando, se o evento  $m$  ocorre, há um evento-base  $f$  que o ocasiona; se  $m^*$  ocorre, há um  $f^*$  que o ocasiona. A nossa intuição inicial quanto ao poder causal de  $m$  em relação a  $m^*$  mostra-se dificultosa em dois momentos: (a)  $m$  como causa de  $m^*$  parece supérfluo, dada a suficiência do evento  $f^*$  para  $m^*$  e a necessidade de um evento físico na ocasião de  $m^*$ ; (2) por isso, se  $m$  tem poder causal sobre  $m^*$ , tal se deve ao poder causal de  $m$  sobre  $f^*$ , que por sua vez é suficiente para  $m^*$ . Mas se  $f$  é suficiente para  $m$  e  $m$  é suficiente para  $f^*$ , então  $f$  é suficiente para  $f^*$ . Isso – lembrar-nos-á o leitor atento – não parece o bastante para inferirmos que  $f$  causa  $f^*$ , isto é, que a seta em negrito expressa uma relação de causação, pois entre  $f$  e  $f^*$  há somente a relação de suficiência, que é ela mesma necessária, mas não suficiente para a relação de causação.<sup>1</sup> No entanto, podemos dizer que há dois eventos *simultâneos e suficientes* para  $f^*$ ,  $f$  e  $m$ . Como, por hipótese,  $m$  causa  $f^*$ , e  $m$  não existiria sem  $f$  (ou sem outro evento físico naquela ocasião), parecemos plausível que  $f$  também cause  $f^*$ . Seria absolutamente anti-intuitivo aceitarmos por hipótese que  $m$  cause  $f^*$  e negarmos que  $f$  cause  $f^*$ , pois sem  $f$  (ou algum outro evento físico),  $m$  não existiria, e  $f$  e  $m$  são contemporâneos. Além disso, se tratamos a noção de causação em termos

<sup>1</sup> Se suficiência fosse suficiente para causação,  $f$  causaria  $m$ , o que não nos parece um resultado intuitivo. Portanto, precisamos de características adicionais à suficiência para termos causação. Uma dessas características é a *prioridade temporal*;  $f$  e  $m$  são simultâneos enquanto que  $f$  e  $f^*$  não são,  $f$  precedendo  $f^*$ . Assim,  $f$  e  $m$  causam  $f^*$ , mas  $f$  não causa  $m$ .



contrafactuais, também temos razões para acreditar que  $f$  causa  $f^*$ , pois se  $f$  não ocorresse,  $m$  não ocorreria, e por hipótese se  $m$  não ocorresse,  $f^*$  não ocorreria. Consequentemente, se  $f$  não ocorresse,  $f^*$  não ocorreria, aceitando-se aqui a transitividade dos contrafactuais.<sup>2</sup> A moral da história é que causação descendente parece gerar uma situação de *superabundância causal*, onde ambos  $f$  e  $m$  causam  $f^*$ .

Voltando ao exemplo da sede e do desejo de saciá-la, o que o modelo de superveniência mente-corpo nos mostra é que a relação de causalidade que supúnhamos existir entre aquela sede e aquele desejo desemboca na situação em que a ativação neural que embasa o desejo é causada pela sede e pela ativação neural que embasa a sede. Portanto, todo evento mental (em nosso exemplo  $m^*$ ) seria ocasionado por um evento físico ( $f^*$ ), que seria causado tanto por um evento mental ( $m$ ) quanto por um físico ( $f$ ). Se nos lembramos, todavia, da tese do fechamento causal do mundo físico (*viz.*, todo evento físico tem pelo menos uma causa física), torna-se obscuro o papel causal de eventos mentais, pois com ou sem eles os eventos físicos – e por conseguinte os mentais que por eles são ocasionados – continuarão ocorrendo. Assim,

(8) As relações causais  $m$ - $a$ - $m^*$  e  $m$ - $a$ - $f^*$  são aparentes, derivando de um (único) processo causal genuíno entre eventos  $f$  e  $f^*$ .

Assim,

(9) Se a superveniência mente-corpo falha, a causação mental é ininteligível; se a superveniência vale, a causação mental também é ininteligível. Logo, a causação mental é ininteligível.

Parece restar-nos o epifenomenismo ou o reducionismo, no nível dos eventos. Vários aceitam o reducionismo, porém asseveram que isso não é o bastante para mostrar que o mental é impotente causalmente em um mundo físico. Davidson (1970), por exemplo, aceita um monismo de

---

<sup>2</sup> Um problema análogo ao mencionado na nota anterior surge em relação aos contrafactuais. De novo, parece-nos que a saída seria diferenciar os contrafactuais em termos temporais.

eventos (i.e., eventos mentais são eventos físicos) sem aceitar a redução do mental ao físico e sem aceitar a impotência causal do mental. Para Davidson, a distinção entre a mentalidade e o físico não deve existir no nível dos eventos, mas sim no nível das características desses eventos. Assim, apesar de eventos mentais serem eventos físicos, o mental tem poder causal, pois causalidade é uma relação extensional entre eventos. Um slogan que capturaria de maneira justa a posição davidsoniana seria: o mental tem poder causal no mundo físico porque eventos mentais têm poder causal no mundo físico. E isso para Davidson é absolutamente consistente com a não-redução do mental ao físico, isto é, com a ausência de leis psicofísicas estritas. Independente de uma avaliação mais precisa do monismo anômalo, porém, a questão crucial parece ser a seguinte: dado que o dilema de Kim nos mostra a implausibilidade de uma distinção entre *ocorrências* [tokens] mentais e físicas, podemos inferir daí que características, ou propriedades mentais sejam impotentes causalmente? Em outras palavras, é o argumento da exclusão causal um argumento que nos força a escolher entre o epifenomenismo *de tipo* e o reducionismo *de tipo*? Essa questão configura o núcleo das preocupações de Kim em relação à diferença causal que mentes fazem em um mundo físico.

Se eventos mentais são eventos físicos e propriedades mentais não são propriedades físicas, então todo evento mental tem dois tipos de características: mentais e físicas. A pergunta ulterior que o partidário da exclusão causal deve fazer parece ser a seguinte: *em virtude de que* eventos mentais têm poder causal? Em virtude de suas propriedades mentais ou de suas propriedades físicas? Se examinarmos o dilema exposto, agora tendo em vista a diferença causal que propriedades mentais possam fazer face à diferença causal que propriedades físicas fazem, percebemos o seguinte. Ao diferenciarmos dois níveis de propriedades, mental e físico, e aceitarmos que todo evento físico tem uma causa física, o papel causal das propriedades mentais se torna também supérfluo, pois em situações em que não há propriedades mentais e há propriedades físicas, as físicas continuarão a fazer a diferença causal que fazem, o mesmo não

ocorrendo em situações em que não haja propriedades físicas e haja propriedades mentais. Isso é o mesmo que dizer que eventos mentais causam eventos físicos *qua* físicos, e não *qua* mentais. Isto é, eventos mentais causam eventos físicos em virtude de suas características físicas e não em virtude de suas características mentais. Portanto, o reducionismo de ocorrências por si só não resolve o problema da relevância causal de mentes em um mundo físico.

Se imaginamos, por exemplo, uma situação em que não haja mentalidade mas haja o mundo físico tal qual o conhecemos, ativações neurais continuariam causando ativações neurais, o mesmo não acontecendo em mundos onde haja mentalidade somente. Essa aparente assimetria entre os domínios mental e físico (i.e., o fato de propriedades mentais não poderem “flutuar desancoradas ao físico”), parece essencial dentro de uma concepção fisicalista da mente; se o mental não dependesse do físico e o físico não fosse independente do mental, não estaríamos dentro dos limites do fisicalismo! A possibilidade de causação mental sem bases físicas *poderia* constituir uma situação, porém não uma situação para um fisicalista. São situações lógicas e metafisicamente possíveis, porém por demais distantes da nossa situação, ou de situações possíveis mas similares à nossa, falhando deste modo em oferecer-nos clarificações interessantes às nossas intuições com relação ao *nosso* problema mente-corpo. Destarte, a pressão para aceitarmos a exclusão no nível das propriedades é forte e não pode ser aliviada senão por uma identificação de propriedades mentais a físicas.

A possibilidade de um epifenomenismo dessas propriedades parece acarretar vários problemas em relação à realidade do mental, dentre os quais: que espécie de existência teria algo que não fizesse diferença causal na rede de eventos que compõe o mundo físico? Não seria um epifenômeno algo absolutamente supérfluo? (Kim (1998)). No entanto, caso aceitemos a identificação de propriedades mentais a propriedades físicas (*viz.*, um fisicalismo de tipo), como lidar com a possibilidade da múltipla realização do mental, intuição essencial do

funcionalismo e precisamente aquilo que originalmente desvitalizou as propostas fisicalistas de tipo em fins da década de cinquenta e início da década de sessenta? Antes de esboçarmos uma resposta a essa pergunta, vejamos como o dilema de Kim se relaciona com o realizacionismo físico.

## 2. O REALIZACIONISMO FÍSICO

A noção de realização em filosofia da mente parece surgir propriamente dentro do contexto das posturas funcionalistas da mente. O funcionalismo, como mencionado, tem como uma de suas teses fundamentais a afirmação de que propriedades mentais são propriedades funcionais, propriedades especificadas de acordo com o papel causal-funcional de seus realizadores, que no caso do funcionalismo fisicalista seriam estados de tipo físico. Assim, parece-nos plausível aceitar que o fisicalismo de tipo e o funcionalismo fisicalista sejam posturas mutuamente excludentes diante do problema mente-corpo, pois usam princípios diversos para a individuação de propriedades mentais. Veremos que isso, entretanto, é um equívoco.

Uma pergunta que devemos responder para que entendamos mais precisamente os compromissos metafísicos do funcionalismo é a seguinte: o que querem dizer os funcionalistas quando asseveram que uma propriedade não-mental qualquer *realiza* uma propriedade mental? Infelizmente, por algum tempo não parece ter havido uma resposta clara para esta pergunta (Kim (1998), p. 7). Aparentemente, a inspiração para o uso desse conceito veio da idéia de que máquinas computacionais caracterizadas matematicamente seriam realizadas, ou implementadas, em dispositivos físicos diversos (*Ibid.*). Assim como uma máquina de Turing poderia, por exemplo, ser realizada em um dispositivo eletrônico, a mente seria realizada no cérebro. No entanto, essas intuições não espelham um esforço especialmente notável para a explicitação da metafísica do funcionalismo.

Ned Block (1990) sugere uma maneira mais precisa de se entender propriedades funcionais. Propriedades funcionais são propriedades de segunda-ordem, definidas a partir de outras propriedades, de primeira-ordem. Em termos gerais, uma propriedade de segunda-ordem é uma propriedade especificada a partir de uma outra propriedade, de primeira-ordem, tal que esta última satisfaça uma condição determinada. Assim, a propriedade ‘... está no estado mental de tipo dor’, por exemplo, seria uma propriedade de segunda-ordem definida a partir de uma propriedade interna do sistema em questão, tal que esta satisfizesse a condição de ter *inputs* e *outputs* característicos da dor. Esta tese, em apresentação mais precisa, equivale ao seguinte:

Seja  $\Gamma$  um conjunto não-vazio de propriedades de primeira-ordem.  $M$  é uma propriedade de segunda-ordem definida sobre o conjunto  $\Gamma$  se, e somente se,  $M$  é a propriedade de ter uma propriedade  $P$ , tal que  $P \in \Gamma$ , e  $C(P)$ , onde  $C$  especifica uma condição sobre os membros de  $\Gamma$  (Kim (1998), p. 20).

As propriedades que pertencem ao conjunto  $\Gamma$  e que satisfazem a condição  $C$  são propriedades dos realizadores da propriedade de segunda-ordem,  $M$ ; são propriedades físicas que sistemas cognitivos têm.

As propriedades funcionais, de acordo com a leitura metafísica exposta na definição acima, são casos especiais de propriedades de segunda-ordem, casos nos quais a especificação  $C$  é uma relação causal vigente entre estados particulares que têm propriedades de primeira-ordem, fixadas as leis naturais.

O que se ganha com essa apresentação metafísica das propriedades funcionais/mentais? A resposta parece clara e consistente com o projeto funcionalista: ganha-se, em primeiro lugar, a abstração do tipo de evento físico, ou da propriedade física, envolvido na relação nômico-causal definidora da propriedade funcional/mental. Para que o funcionalista mostre como se individualizam propriedades mentais basta a menção do papel causal da propriedade interna em questão, independente desse

papel ser desempenhado por uma ativação neural ou por um estado hidráulico do sistema cognitivo relevante.

Braddon-Mitchell e Jackson (1996) sublinham que o funcionalismo é compatível com duas posições metafísicas: (a) o estado mental (e.g. 'dor') é o estado realizador (de primeira-ordem, e.g., 'ativação da fibra C'); (b) o estado mental é o estado funcional (de segunda-ordem). A posição (a) garante poder causal ao estado mental no sentido (b). Embora defendam esta posição, eles rejeitam, contudo, o essencialismo da teoria fisicalista de tipo: o que um estado realizador (de primeira-ordem) é capaz de "fazer" não deve ser visto como uma propriedade essencial (intrínseca) desse estado. Nesse sentido, funções (enquanto propriedades de segunda-ordem) não são essenciais aos estados físicos que as realizam: diferentes estados físicos podem desempenhar a mesma função. Ou ainda, o mesmo estado físico pode desempenhar uma função numa (num tipo de) estrutura e não desempenhá-la em outra estrutura. Um coração mecânico, por exemplo, pode desempenhar a mesma função que os nossos corações-músculo. As fibras C e D podem, ambas, desempenhar a função de dor (especificada abstratamente em termos de um papel causal) em humanos e em golfinhos, respectivamente. Não obstante, lembramos que a pergunta que Kim coloca ao funcionalista continua sem resposta: dado esse entendimento geral do mental, será que o fato de *ocorrências* mentais causarem *ocorrências* físicas é razão suficiente para concluirmos que *tipos* mentais façam diferença causal em um mundo físico?

Diante da interpretação de propriedades funcionais como casos especiais de propriedades de segunda-ordem há algumas qualificações a serem feitas, para que enxerguemos com maior nitidez os seus corolários. Kim, em seus últimos trabalhos (e.g., (1998)), expressou com clareza tais qualificações.

Em primeiro lugar, se temos dois sistemas com estruturas funcionais diferentes (i.e., se a rede causal-funcional que liga os inputs aos *outputs* dos dois sistemas for diferente), há variação nos realizadores. Com efeito, não se pode garantir que um mesmo realizador terá o

papel causal típico de dor em dois organismos que tenham organizações neurais e motoras diferentes. Porém, desde que a estrutura causal-funcional dos organismos em questão seja fixada, podemos dizer *ceteris paribus* que se os dois estiverem no mesmo estado interno, estarão no mesmo estado funcional.

Em segundo lugar, se os dois sistemas estão inseridos em mundos diferentes,  $w$  e  $w'$ , tal que as leis físicas sejam relevantemente diferentes em  $w$  e  $w'$ , então não se pode dizer em geral que, dado um mesmo estado interno e uma mesma estrutura cognitiva, os sistemas estarão no mesmo estado mental. Inversamente, se as leis dos mundos nos quais os dois sistemas estejam inseridos são as mesmas, podemos dizer que se os sistemas partilham da mesma estrutura causal-funcional, um mesmo estado interno realiza neles o mesmo estado funcional. Essa dependência na constância das leis naturais dos mundos nos quais os sistemas relevantes se inserem sugere-nos a íntima relação entre o poder causal de um estado interno e as leis físicas que os limitam. Desse modo, se as leis são diferentes, os poderes causais também o serão.<sup>3</sup>

Essas duas qualificações mostram-nos um resultado interessante. Se dois sistemas têm a mesma estrutura causal-funcional, então dado que estejam em um mesmo tipo de estado interno, estarão *necessariamente* no mesmo tipo de estado funcional, onde a necessidade em questão é *nomológica*. Assim, a tese do realizacionismo físico acarreta a superveniência nomológica mente-corpo, pois vale a covariância estabelecida

---

<sup>3</sup> É interessante notar que se endossamos um realismo em relação a leis, em particular às leis causais, tal intimidade deixa de se mostrar com toda a sua força. Além disso, se endossamos um realismo em relação às relações causais, tal intimidade se enfraquece mais ainda. No entanto, não acreditamos que o núcleo do argumento de Kim seja alterado caso escolhamos uma posição realista em relação à natureza de leis e das relações causais. Para uma clara apresentação das diferenças entre posições realistas e reducionistas no que diz respeito a leis e relações causais vale examinar Tooley (1990).

pela superveniência somente no caso em que os sistemas ocorram em mundos com as mesmas leis físicas.

O que ganhamos com isso, entre outras coisas, é a explicação do porquê haver superveniência entre propriedades mentais e propriedades físicas. Como tem insistido Kim (e.g., (1990), (1997), (1998)), a relação de *covariância* por si só não seria capaz de gerar – pelo menos não trivialmente – uma relação de *dependência* entre a mente e o corpo, pois não há nada na covariância que assegure a assimetria necessária para fundarmos a dependência do mental ao físico. E se não há a dependência, não nos parece muito claro como poderíamos explicar o nível superveniente em termos do nível subveniente. Uma saída seria tomar a relação de dependência como primitiva (cf., Kim (1990)). Já no caso da relação de realização a questão é diferente, pois geralmente aceita-se que esta seja, de fato, uma relação de dependência (Kim (1992)). Se uma propriedade *M* é realizada por uma propriedade *F* em certas espécies de organismos, então a exemplificação de *M* *depende* da exemplificação de *F*, dada a constância nomológica mencionada.

Assim, essa maneira de individuar propriedades mentais postula a existência de propriedades de primeira-ordem, nesse caso físicas, sendo isso aparentemente o bastante para afirmarmos que propriedades funcionais dependam ontologicamente de propriedades físicas. No caso inverso, a analogia não vale, pois para que os realizadores existam não é preciso que exista aquilo que eles realizam. Desse modo, parece haver uma *primazia ontológica* dos realizadores, sendo precisamente isso o que gera o padrão superficial de covariância entre as propriedades físicas dos realizadores e as propriedades funcionais, padrão esse exemplificado e exibido pela relação de superveniência.

### 3. O PROBLEMA DA CAUSAÇÃO MENTAL E O REALIZACIONISMO FÍSICO

Imaginemos que um evento mental, *m*<sup>\*</sup>, (e.g., a dor de cabeça de João Silva, às 16:30, do dia 15 de Outubro de 2001) seja causado pelo



evento mental  $m$  (e.g., a angústia de João, um pouco antes daquela ocasião). Nesse caso, a dor de cabeça,  $m^*$ , é realizada por um evento físico tal que este seja suficientes para  $m^*$ . Em outras palavras, a dor de cabeça é aquele evento físico, pois “ser dor de cabeça” é uma característica de certos eventos físicos. Em outras palavras, um evento ser uma dor de cabeça é, consoante o realizacionismo físico, ele ser um evento físico causado por eventos-input e eventos-output relevantes. Assim, não precisamos do argumento da exclusão para desvelar a implausibilidade da diferença entre eventos, ou *ocorrências*, mentais e físicas. O realizacionismo físico acarreta a identidade de ocorrências psicofísicas.

Isso, no entanto, não quer dizer que propriedades mentais sejam propriedades físicas, pois identidade de *ocorrências* não acarreta identidade de tipos, ou propriedades. Como lidar com essa situação? Kim propõe que aceitemos um princípio chamado ‘princípio da herança causal’. Reza esse princípio que o poder causal de exemplificações de propriedades de segunda-ordem é herdado do poder causal de exemplificações das propriedades de primeira-ordem que realizam as de segunda-ordem. Isto é, o poder causal de um evento mental é o mesmo que o de um evento físico, o que sugere que propriedades mentais não fazem diferença causal. Eventos mentais causariam eventos físicos em virtude de suas características físicas e não de suas características mentais, de acordo com o princípio da herança causal. De novo, voltamos ao impasse anterior: o epifenomenismo das características, ou propriedades, mentais ou o reducionismo de propriedades mentais a propriedades físicas. Se aceitamos o *Dictum de Alexander* (*viz.*, ser real é ter poder causal), resta-nos a redução. Se não o aceitamos, podemos ainda permanecer dualistas, porém epifenomenistas.

Mas como – poder-se-ia indagar – podemos responsabilmente optar pela redução se esta é incompatível com a múltipla realização do mental, intuição que goza de unanimidade no debate mente-corpo? A proposta de Kim é a identificação de propriedades mentais a

propriedades físicas em estruturas específicas. Assim, nessas estruturas, ter mente é ter aquilo que realiza a mente, o mesmo não podendo ser afirmado irrestritamente.

A delimitação dessas estruturas é, sem dúvida alguma, tópico de controvérsia. No entanto, a resolução disso seria produto de um trabalho empírico sobre a cognição. Destarte, somente a pesquisa empírica poderia nos fornecer os subsídios para talharmos as identidades específicas.

Mas e a linguagem do mental? Eliminamo-la? Se a resposta for positiva, a posição de Kim se assemelha ao materialismo eliminativo de Chuchland (1981). Se a resposta for negativa, a pesquisa científica sobre a cognição em termos irrestritos se mostra gravemente ameaçada. Poder-se-ia ter *ciências* cognitivas e não *uma* ciência cognitiva em geral.

#### 4. O FUNCIONALISMO CONCEITUAL

Como conseqüência do problema da exclusão causal, em seus últimos trabalhos Kim defende que o pacote [dualismo de propriedades + fisicalismo] não é sustentável e propõe o que poderíamos chamar de um “funcionalismo conceitual”.<sup>4</sup>

Segundo Kim, quando dizemos que a ‘dor’ em geral é ocasionada por uma propriedade física de primeira-ordem, tal que esta última satisfaça o papel causal-funcional que define a dor, estamos ilicitamente tratando *conceitos* mentais, definidos funcionalmente, como se fossem *propriedades* mentais. Por outro lado, quando tratamos de espécies ou estruturas específicas, podemos falar de propriedades mentais e temos a redução direta do mental ao físico. No ser humano, ‘dor’ não é nada mais do que a ativação da fibra-C, sendo uma outra propriedade no caso dos moluscos, ainda outra no caso dos marcianos, e assim por diante. O resultado dessa identificação é a manutenção do poder causal das

---

<sup>4</sup> Kim mantém-se não-reducionista com respeito à consciência, aos *qualia*, contudo.

propriedades mentais restritas, manobra esta que nos salva a realidade do mental.

Um conceito mental  $M$  é definido da seguinte forma:

O conceito de  $M \approx$  o conceito de ter alguma propriedade de primeira-ordem  $\phi$  tal que estar em  $I$  causa um indivíduo ter  $\phi$  e ter  $\phi$  causa um indivíduo ir para  $O$ .<sup>5</sup>

As relações de superveniência e de dependência não se aplicariam mais (de modo não-trivial) a *propriedades* irreduzíveis, mas teriam um novo uso “enquanto relações entre conceitos e/ou entre expressões lingüísticas...” (Horgan (1997), p. 178).

Já no nível ontológico, eliminar-se-iam características mentais irrestritas em virtude de sua heterogeneidade causal e identificar-se-iam características mentais a características físicas, restritamente.

Outra característica interessante da nova postura de Kim é que a funcionalização (i.e., a leitura de propriedades de alto-nível em termos de propriedades funcionais) passa a ser vista como um modelo de redução (que superaria as dificuldades do modelo nageliano de redução).<sup>6</sup>

## 5. A ANÁLISE CAUSAL DOS CONCEITOS MENTAIS DE ARMSTRONG E LEWIS

O funcionalismo conceitual para o qual Kim parece pender foi antecipado, de certo modo, por Armstrong e por Lewis em trabalhos independentes.

Já no seu livro *A Materialist Theory of Mind* (1968), Armstrong propõe-se a contribuir para uma “análise causal” de conceitos mentais. Armstrong defende que a identificação entre estados mentais e estados físicos proposta por Smart e Place deve, a rigor, ser precedida por um

---

<sup>5</sup> Caracterizar precisamente o status da relação ‘ $\approx$ ’ é um dos problemas, segundo David (1997), que necessitam ser investigados pelos defensores desse tipo de funcionalismo conceitual.

<sup>6</sup> Braddon-Mitchell e Jackson (1996) defendem, por exemplo, que a melhor maneira de introduzir a teoria da identidade é via o funcionalismo!

trabalho filosófico de análise dos conceitos mentais. Armstrong via essa análise como sendo, na verdade, independente de uma solução particular para o problema mente-corpo, como a proposta pela teoria da identidade e, mesmo, de forma mais geral, independente do materialismo, ou fisicalismo. Mas essa análise pode “abrir caminho”, segundo ele, para uma solução fisicalista ((1991), p. 183).

Um conceito mental, segundo a análise que propõe Armstrong, é o conceito de um estado físico que desempenha um papel causal numa cadeia envolvendo estímulos, respostas comportamentais e, via de regra, também outros estados ‘internos’:

... o conceito de um estado mental envolve essencialmente, e é esgotado pelo conceito de um estado que é apto a ser a causa de certos efeitos ou apto a ser o efeito de certas causas ((1991), 183/1).<sup>7</sup>

É importante ressaltar que, em seu livro, Armstrong não fala de propriedades mentais, mas somente de conceitos mentais. Ele não compartilha de um dualismo de propriedades como o associado ao funcionalismo clássico (“hard-core”).<sup>8</sup>

Armstrong distingue, portanto, respostas científicas para o problema da natureza intrínseca dos estados mentais, de análises filosóficas dos conceitos mentais, que deveriam preceder ou que são pressupostas por aquelas respostas.

---

<sup>7</sup> Grande parte do livro *A Materialist Theory of Mind* (1968) é dedicado a propor “análises causais” de diversos conceitos mentais, como os de percepção, crença, propósito, etc.

<sup>8</sup> O funcionalismo só veio a tornar-se uma posição respeitável em filosofia da mente, posteriormente. Os primeiros artigos do Putnam-funcionalista datam de 1966-7. De toda forma, a ontologia associada ao funcionalismo (a distinção entre propriedades de primeira-ordem e de segunda-ordem) não era clara à época, nem para os próprios funcionalistas!

Smart e Place só teriam tratado da segunda parte de um “programa”<sup>9</sup>, com um caráter mais propriamente científico, e não da primeira parte, a análise conceitual, que possui um caráter filosófico.<sup>10</sup> Esta primeira parte deve ser “plausível independentemente” (Armstrong (1991), p. 183) da segunda<sup>11</sup>:

Notar-se-á que o argumento tem dois ‘movimentos’ distintos. Como um primeiro passo, propõe-se uma certa análise lógica dos conceitos mentais. Esta é uma tese conceitual. Ela deve ser estabelecida ou refutada da mesma maneira (qualquer que seja ela) que é estabelecida ou refutada qualquer outra peça de análise expressamente lógica. Ela pode talvez ser chamada de análise Causal dos conceitos mentais. Dela não se segue o Materialismo, mas tampouco o exclui. O segundo passo do argumento é o de identificar esses estados internos com estados físico-químicos do cérebro. Essa é uma identificação contingente ou científica, e ela dá lugar ao Materialismo de estado-Central [*Central-state Materialism*] (Armstrong (1968), p. 91).

A situação seria análoga às identificações teóricas<sup>12</sup> nas ciências, como a de Gene  $\equiv$  DNA. O conceito de gene é definido/analísado relacionalmente, incluindo referências à observação (traços fenotípicos,

---

<sup>9</sup> Armstrong ((1991), 184/2) refere-se explicitamente a Lakatos quando diz que não se trata de uma “teoria causal” mas propriamente de um “programa”. O seu livro de 1968 inicia esse programa, que pretende definir todos os conceitos mentais através de análises causais.

<sup>10</sup> No artigo de 1977/1991, a metafilosofia de Armstrong é claramente naturalista, defendendo ele que a análise depende, na verdade, de comprometimentos ontológicos por parte do filósofo, que deve se envolver com “questões de primeira-ordem”, contrariando a meta-filosofia de Wittgenstein e de Ryle, que pretenderam limitar o trabalho filosófico à análise conceitual: “... a análise de conceitos é um meio pelo qual o filósofo dá a sua contribuição para as grandes questões gerais, não acerca de conceitos, mas acerca de coisas” (Armstrong (1991), p. 182).

<sup>11</sup> Sobre essas duas etapas, a de definição (análise) filosófica e a de descoberta científica, ver também Lewis (1991a), p. 204.

<sup>12</sup> A identificação é ‘teórica’ porque DNA é um termo teórico e, não, observacional. Do mesmo modo, os termos mentais são teóricos.

por exemplo). Em seguida, uma investigação científica pode identificar a molécula que desempenha o papel causal especificado pelo conceito de gene.

Lewis apresenta esta proposta como a de "... uma teoria materialista da mente que associa teses da identidade psicofísica tipo-tipo [*type-type*] com uma maneira behaviorista ou funcionalista de caracterizar estados mentais como o de dor". E continua mais adiante:

A nossa visão é que o conceito de dor, ou mesmo de qualquer outra experiência ou estado mental, é o conceito de um estado que desempenha um certo papel causal, um estado com certas causas e efeitos típicos. É o conceito de um estado apto a ser causado por certos estímulos e apto a causar um certo comportamento. É o conceito de um membro de um sistema de estados que, juntos, realizam mais ou menos o padrão [*pattern*] de generalizações causais descritas pela psicologia de senso-comum (Lewis (1991b), p. 230).

Os conceitos (mentais) e os termos/predicados mentais são considerados não-rígidos. O conceito de dor não é o conceito do estado neural que realiza a dor num indivíduo ou numa "população" particular. O conceito de 'dor' pode designar diferentes estados físicos/neurais em diferentes indivíduos ou em diferentes populações ou espécies. É uma questão contingente que a dor em um indivíduo ou população seja idêntica a um particular estado neural.<sup>13</sup> Temos somente um reducionismo local. Ou melhor, é uma questão contingente que um tipo particular de propriedade física (e.g., a ativação da fibra *C*) desempenhe a função causal especificada pelo conceito de dor.

O funcionalismo conceitual de Armstrong-Lewis difere do de Kim, contudo, porque os primeiros voltavam-se para análises de conceitos mentais em *explicações* psicológicas de senso-comum, enquanto

---

<sup>13</sup> Horgan (1997) classifica a posição de Armstrong-Lewis como um "materialismo monista de redução local" (MMRL).

que o último sempre esteve preocupado com a *metafísica* das relações entre os âmbitos do mental e do físico.

A análise de Armstrong-Lewis permite resolver o problema da exclusão causal (que não era, contudo, o problema deles à época) identificando a propriedade mental com a propriedade física que desempenha, num indivíduo (numa espécie, população, ou tipo de estrutura) o papel causal especificado.

No que diz respeito à múltipla realização, não parece ser admissível, para Kim, a definição de propriedades de segunda-ordem em termos de uma disjunção de propriedades de primeira-ordem. Vários problemas parecem se colocar diante dessa tentativa de solução (Kim (1990)), problemas que põem em dúvida a própria capacidade nomológica e, portanto, a possibilidade de uma ciência irrestrita do mental (Kim (1992b)). O que há é a múltipla *aplicação* do conceito mental e não a múltipla *realização* de propriedades mentais.

David (1997) frisa que conceitos mentais não são uma questão de ontologia, mas de epistemologia e de interesses: “Concepts reflect our attempts to gain epistemic access to what is out there and our need to communicate about what is out there”.

Ou seja, há uma dimensão pragmática inescapável na aplicação dos (nossos) conceitos mentais: “... conceitos de estado mental não-restritos... cortam a natureza [*carve nature*] mais de acordo com as nossas necessidades do que nas suas juntas [*along its joints*]” (*Ibid.* 137).

Papéis funcionais são definidos em termos de acesso epistêmico, de evidência disponível publicamente, em outras palavras, em termos do modo como interagimos com o mundo; os conceitos funcionais traduzem esses aspectos epistemológicos e pragmáticos e, portanto, não designam espécies naturais. Eles especificam, em vez disso, as condições para a aplicação de um conceito ou do predicado correspondente a algo; no caso da definição dos conceitos mentais, os ‘inputs’ e ‘outputs’ são públicos (David, *ibid.* 138).

Block e Stalnaker, por sua vez, contestaram em artigo recente (1999) a análise conceitual de Armstrong-Lewis e o modelo funcional de redução (que Kim está disposto a aceitar, em função das dificuldades colocadas pelo fisicalismo não-redutivo). Eles defendem que as identidades estado mental  $\equiv$  estado físico são identidades *a posteriori*, identidades kripkeanas, como água  $\equiv$  H<sub>2</sub>O. Não há necessidade de uma análise conceitual *a priori* do conceito de 'água', ou, de maneira análoga, do conceito de 'dor' para que se possa estabelecer tais identidades. As identidades kripkeanas seriam justificadas pelo sucesso obtido em nossas explicações científicas. Esta posição parece desembocar, contudo, diretamente no reducionismo, mesmo que restrito.

Um outro ponto interessante que notamos *en passant* em relação a essas identidades é o seguinte: a posição reducionista de Kim parece responder às críticas de Kripke (1980) à contingência das identidades mente-corpo. Kripke aparentemente tinha em mente a posição de Smart (1959) quando criticou a identidade mente-corpo, segundo a qual tipos mentais são tipos físicos, contingentemente. Para Kripke, como o predicado 'dor' designa rigidamente o *quale*, e, além disso, 'ativação da fibra-C' também designa rigidamente aquele tipo de ativação neural, a identidade há de ser necessária, mesmo que *a posteriori*. Kim, ao identificar propriedades mentais a propriedades físicas em estruturas específicas, dada a fixidez das leis naturais, oferece-nos identidades necessárias, onde a necessidade é entendida nomologicamente. Assim, a análise de Kim tem a vantagem de consoar com as intuições de Kripke no que diz respeito às identidades psicofísicas.

## 6. HÁ AINDA LUGAR PARA UM FISCALISMO NÃO-REDUTIVO?

Diante disso tudo, restam esperanças para os fisicalistas não-redutivos? O pacote {funcionalismo conceitual + monismo materialista} é o único possível para os fisicalistas, diante dos dilemas salientados por Kim? Há quem resista à aparente plausibilidade dos argumentos de Kim.



Horgan (1997) continua defendendo o que ele chama de um “compatibilismo causal robusto” (PCR), na direção da defesa que faz Fodor de espécies naturais genuínas, com poder causal autônomo, nas ciências especiais.

Horgan defende que só há duas posições filosoficamente estáveis para o realista com respeito ao mental<sup>14</sup>: um “incompatibilismo mental” ou um “compatibilismo causal robusto”. Não haveria lugar para posições conciliatórias:

Não há na paisagem conceitual qualquer base estável para um compatibilismo medroso [*angst-ridden*] que tente simultaneamente abraçar um realismo mental, repudiar identidades psicofísicas tipo-tipo e acomodar o espírito do argumento de exclusão causal (Horgan (1997), p. 181).

As duas posições “estáveis” são caracterizadas da seguinte maneira:

(1) “Incompatibilismo mental”: as propriedades mentais não têm poder causal autônomo, só as físicas; a esta tese associa-se uma identidade de tipo. Este é o “pacote” a que nos referimos há pouco.

(2) “Compatibilismo causal robusto”: há genuínas espécies naturais nas ciências especiais e elas têm poder causal autônomo. Esta é a posição defendida por Horgan.

Os defensores desta segunda posição terão, contudo, que explicar porque o argumento da exclusão causal parece intuitivamente aceitável (ou localizar os erros “sutis” neste argumento). Horgan sugere uma explicação cognitiva para isso que não nos parece muito promissora ((1997), p. 179). A seguir apresentaremos uma outra posição que se pretende não-reduziva.

---

<sup>14</sup> Eliminativistas como os Churchland são ‘não-realistas’ com respeito ao mental.

## 7. O FUNCIONALISMO ESTRUTURAL DE M. DAVID

David (1997) distingue o gênero de funcionalismo proposto recentemente por Kim – um funcionalismo conceptualizado, *FC* – e o funcionalismo, digamos, ortodoxo (“hard-core funcionalism” - *HCF*). David defende uma alternativa a esses dois gêneros de funcionalismo, um funcionalismo estrutural (*hard core structural functionalism*<sup>15</sup>) diante dos argumentos de Kim.

Como vimos, o *HCF* define um estado mental de tipo *M* (ou uma propriedade mental *M*) da seguinte forma:

A propriedade  $M = a$  propriedade de ter alguma propriedade de primeira-ordem  $\phi$  tal que estar em *I* causa um indivíduo ter  $\phi$  e ter  $\phi$  causa um indivíduo ir para *O*.

Simbolicamente temos:

$$(C) \quad x \text{ tem } M \Leftrightarrow (\exists \phi) [\phi x \ \& \ (I \rightarrow \phi \rightarrow O)].$$

No *HCF*, o estado mental de tipo *M* é identificado ao estado funcional.<sup>16</sup>

Kim rejeita o *HCF* em seus vários pontos. Os estados mentais seriam restritos a espécies ou a tipos de estrutura.

<sup>15</sup> Ainda “hard-core” porque não considera inevitável a saída “conceitual”.

<sup>16</sup> O *HCF* não pode, portanto, ser visto como um fisicalismo reducionista, já que “... os estados funcionais em questão são constituídos de estados físicos juntamente com propriedades e relações chamadas ‘tópico-neutras’ [Smart] (...) O resultado de tudo isso é que o funcionalismo não é um reducionismo fisicalista ‘hard-core’. Ele reduz estados mentais a estados físicos-junto-com-tópico-neutros [*physical-cum-topic-neutral states*]. Esse tipo de fisicalismo enfraquecido [*weak-minded*] deve ser suficiente para funcionalistas ‘hard-core’. Eles assinalarão que aquelas propriedades tópico-neutras, embora não puramente físicas, não são tampouco, no final das contas, propriedades mentais: não existe qualquer base para um dualismo aqui” (David (1997), p. 135).

Vimos que Kim reinterpreta (C) acima como definindo, não uma propriedade, mas um conceito. O funcionalismo de Kim pode ser definido como:

*FK*: O conceito de  $M \approx$  o conceito de ter alguma propriedade de primeira-ordem  $\phi$  tal que estar em  $I$  causa um indivíduo ter  $\phi$  e ter  $\phi$  causa um indivíduo ir para  $O$ .<sup>17</sup>

David defende, surpreendentemente, que o problema central no funcionalismo, tanto o *FK* quanto o *HCF*, seria a sua incompatibilidade com a múltipla realização! Para demonstrar isso, David pressupõe que asserções a respeito da causação de propriedades são equivalentes ou implicam generalizações universais nomologicamente necessárias. Obtemos, desse modo, uma versão expandida de (C):

$$(C1) \quad x \text{ tem } M \Leftrightarrow (\exists \phi) [\phi x \ \& \ \forall y (Iy \rightarrow \phi y \rightarrow Oy)].$$

(C1) não admitiria a múltipla realização porque os quantificadores estão invertidos. Segundo esse enunciado, para todos os organismos ( $y$ ), existe uma propriedade física  $\phi$  que desempenha o papel funcional característico do estado mental  $M$ :

Segue-se de (C1) que o papel funcional não pode ser desempenhado por propriedades físicas diferentes em organismos diferentes (David, *ibid.*, 145).

O diagnóstico de David é que o problema estaria na “ramseificação”, um procedimento para eliminar os termos teóricos de uma

---

<sup>17</sup> Caracterizar precisamente o status da relação ‘ $\approx$ ’ é um dos problemas, segundo David, que necessitam ser investigados pelos defensores desse tipo de funcionalismo conceitual.

teoria substituindo-os por variáveis predicacionais quantificadas existencialmente.

A proposta de David para escapar desse impasse é descartar a abordagem de Ramsey, de modo a inverter a posição dos quantificadores:

... precisamos dizer que estar com dor é ser um desses organismos nos quais alguma propriedade física  $\phi$  ocupa o papel característico e ter  $\phi$  (*Ibid.*, :145-6).

David acredita que um “funcionalismo estrutural” resiste ao argumento da herança causal, que fora desastroso para o *HCF* e, portanto, resiste à conceptualização do funcionalismo. A idéia básica é que uma propriedade física tem um certo poder causal *porque* o organismo possui um tipo de estrutura física. Este tipo de estrutura é que cria as condições para que uma propriedade física desempenhe a função causal que efetivamente desempenha. Simbolicamente, o funcionalismo estrutural define as propriedades mentais da seguinte forma:

$$(C2) \ x \text{ tem } M \Leftrightarrow \{ \exists ST \} \{ STx \ \& \ (\forall y) (STy \rightarrow \exists \phi (\phi x \ \& \ (Iy \rightarrow \phi y \rightarrow Oy))) \}$$

Nas palavras de David:

A causação de  $e$  por  $P1$  é, ele mesma, causada pela ocorrência de  $P1$  dentro de uma estrutura de tipo  $ST1$  (se  $P1$  não tivesse ocorrido dentro de uma estrutura  $ST1$ , ele não teria causado  $e$  nessa ocasião a menos, obviamente, que  $P1$  tivesse ocorrido dentro de uma das outras estruturas que determinam o papel funcional característico de  $M$ ). A idéia seria que os poderes causais de  $M$  derivam das estruturas físicas incorporadas em organismos tendo  $M$ . Essas estruturas físicas funcionam como causas que permitem os realizadores físicos de  $M$  causarem os eventos físicos que eles causam (David, *ibid.*, 146).

O tipo de estrutura é caracterizado pelo mesmo papel funcional/causal desempenhado por diferentes propriedades físicas, daí a

múltipla realização. Esse tipo funcional, digamos assim, é possibilitado pelo tipo de estrutura.

O funcionalismo estrutural proposto por David nos parece, contudo, compatível com o que diz Kim, pois em estruturas específicas o que faz diferença causal é o físico e não o mental enquanto mental.

## CONCLUSÕES

O funcionalismo é uma tentativa de articular imagens de natureza e intuições que nos parecem difíceis de abandonar:

- (1) Uma concepção estratificada da realidade, em que há níveis ontológicos distintos, gozando de relativa autonomia;<sup>18</sup>
- (2) As capacidades e propriedades associadas ao mental necessitam uma abordagem abstrata, propriamente funcional, que aponta para uma ordem ou nível ontológico mais elevado do que o descrito pela física (ou, se quiserem, pela neurofisiologia). Por trás dessa tese está a intuição da múltipla instanciação do mental.

Acreditamos, por exemplo, que as noções de “função”, de “forma” e de “estrutura” ainda podem contribuir para a elucidação da realidade e da especificidade do mental num mundo físico. A metáfora computacional permanece sedutora, em especial a idéia de que mentes são máquinas virtuais que impõem restrições aos processos físico-químicos que ocorrem nos cérebros dos animais (e, possivelmente, em outros tipos de ‘hardware’), explorando a sua plasticidade e potencializando-os.<sup>19</sup>

Ademais, devemos ser consistentes não somente numa postura fisicalista (embora não necessariamente redutiva), mas também natura-

---

<sup>18</sup> Kim (1998), por exemplo, tem algumas reflexões interessantes a respeito da distinção entre níveis e ordens ontológicos.

<sup>19</sup> Esta é uma idéia que se encontra em Dennett, entre outros. Ver Dennett (1991), esp. cap. 7, seção 7.

lista.<sup>20</sup> A proposta de Sober, por exemplo, de tentar aproximar a noção de “função” adotada pelo funcionalismo da noção de “função” adotada em biologia é muito sugestiva. Tentativas de “análise” dos conceitos mentais, como a propugnada por Armstrong entre outros filósofos, em vez de pretenderem manter-se *a priori* – e, portanto, prisioneiras de nossas intuições de senso comum – deveriam interagir com os resultados da pesquisa científica, em suas dimensões teórica e empírica, abrindo caminho para uma eventual revisão daquilo que consideramos concebível através de nossos experimentos de pensamento. É claro que as imagens de natureza/de mente e intuições com as quais estamos comprometidos, e que foram explicitadas acima, também estão sujeitas a revisão!

Não obstante isso, não enxergamos, no momento, *como* uma revisão do argumento da exclusão causal poderia ser levada a cabo. Os argumentos de Kim, simples e gerais, representam um desafio genuíno a um programa fisicalista não-reducionista comprometido com o naturalismo. Deste modo, uma saída para o funcionalista-fisicalista seria a explicitação de uma noção alternativa de propriedade funcional, que não o comprometesse com a superveniência psicofísica, com o realizacionismo físico e, por conseguinte, com a exclusão causal. Todavia, enquanto o funcionalista não dispuser de uma *história positiva* nessa direção, o argumento de Kim parece sugerir que funcionalismo e não-redução são mutuamente excludentes, ao contrário do que se imaginou.

*Abstract: What has placed functionalism in the center of the debate around the mind-body problem in the last three decades seems to have been its capacity to conciliate physicalist intuitions with non-reductionism: if, on the one hand, the functionalist of physicalist persuasion postulates the existence of physical entities only, distributed in a stratified ontological model, on the other, she does not shy away from a realism regarding these entities's properties, in particular with respect to the ones related to sensing and representing. Mind-body supervenience seems to have come along to bring light upon these non-reductionist and physicalist intuitions. Kim, notably in the eighties, suggested supervenience to offer a clear sense in which the relation between mentality and*

---

<sup>20</sup> Embora seja comum equacionar-se fisicalismo, reducionista e naturalismo, isso é um equívoco. Ver Abrantes (1998).

*the physical could be understood. In the last years, nonetheless, he has become increasingly skeptic about the possibility of our understanding mental causation through mind-body supervenience. He sets forth a dilemma, according to which mental causation is unintelligible, even were supervenience true. In Kim's view, then, mind-body supervenience turns out to compose the problem, instead of expressing a solution to it. Moreover, since mind-body supervenience seems to apply to all minimally physicalist positions on the mind-body problem, Kim's misgivings suggest something critical with non-reductive physicalism in general. We shall argue in this paper that, although Kim's misgivings seem justified, it would be headlong to assume the falsity of non-reductive physicalism.*

## BIBLIOGRAFIA

- ABRANTES, P. (1998). "Naturalismo epistemológico: introdução". In: Évora, F.; Abrantes, P. (eds.) *Cadernos de História e Filosofia da Ciência* (CLE/UNICAMP), série 3, v. 8, n. 2, pp. 7-26.
- AMARAL, F. (2001). *Causação Mental e Redução*. Dissertação de Mestrado defendida no Departamento de Filosofia da UnB em 2001 e orientada pelo Prof. Paulo Abrantes.
- \_\_\_\_\_. (2002). Causação Mental: onde estivemos e onde estamos, *Psicologia Teoria e Pesquisa*, v.17, n. 3, pp.235-24.
- ARMSTRONG, D.M. (1968). *A Materialist Theory of Mind* (Londres, Routledge).
- \_\_\_\_\_. (1991). The Causal Theory of Mind. In: Rosenthal, D. (ed.) *The Nature of Mind* (New York, Oxford University Press), pp. 181-8. Publicado originalmente em 1977.
- BLOCK, N. (1990). Can the Mind Change the World? In: Boolos, G. (ed.) *Meaning and Method* (Cambridge, Cambridge University Press).
- BLOCK, N. & STALNAKER, R. (1999). Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review*, v. 108, pp. 1-46.
- BRADDON-MITCHELL, D. & JACKSON, F. (1996). *Philosophy of Mind and Cognition* (Oxford, Blackwell).

- CHURCHLAND, P. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy*, LXXVIII, n. 2, pp. 67-90.
- DAVID, M. (1997). Kim's Functionalism. *Philosophical Perspectives*, v. 11 (Mind, Causation, and World), pp. 133-148.
- DAVIDSON, D. (1970). Mental Events. In: Lawrence Foster & J. W. Swanson (eds.), *Experience and Theory* (Amherst, The University of Massachusetts Press), 1970. Reimpresso em *Essays on Actions and Events* (Oxford, The Clarendon Press), 1980.
- DENNETT, D. (1991). *Consciousness Explained* (Boston, Little, Brown and Company).
- FODOR, J. (1965). Explanations in psychology. In: Black, M. (ed.). *Philosophy in America* (London, Routledge & Kegan Paul).
- HORGAN, T. (1997). Kim on Mental Causation and Causal Exclusion. *Philosophical Perspectives*, v.11 (Mind, Causation, and World), pp. 165-183.
- KIM, J. (1990). Supervenience as a Philosophical Concept. *Metaphilosophy* 21, pp. 1-27.
- \_\_\_\_\_. (1992a). 'Downward Causation' in Emergentism and Non-reductive Physicalism. In: Beckermann, Flohr & Kim (eds.) *Emergence or Reduction?* (Berlim, De Gruyter).
- \_\_\_\_\_. (1992b). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research* 52, pp.1-26.
- \_\_\_\_\_. (1997). The Mind-Body Problem: Taking Stock after Forty Years. *Philosophical Perspectives*, v. 11, pp. 185-207.
- \_\_\_\_\_. (1998). *Mind in a Physical World* (Cambridge (MA), MIT Press).
- LEWIS, D. (1991a). Psychophysical and Theoretical Identifications. In: Rosenthal, D. (ed.) *The Nature of Mind* (New York, Oxford University Press), pp. 204-210. Publicado originalmente em 1972.



- \_\_\_\_\_. (1991b). Mad Pain and Martian Pain. In: Rosenthal, D. (ed.) *The Nature of Mind* (New York, Oxford University Press), pp. 229-235. Publicado originalmente em 1980.
- PUTNAM, H. (1965). Brains and Behavior. In: Butler, R. J. (ed.) *Analytical Philosophy*, Second Series (Oxford, Blackwell).
- \_\_\_\_\_. (1967) The Nature of Mental States. In: Capitan & Merrill (eds.) *Art, Mind and Religion* (Pittsburgh, University of Pittsburgh Press).
- SMART, J. (1959). Sensations and Brain Processes. *Philosophical Review* 68, pp. 141-156.
- TOOLEY, M. (1990). Causation: Reductionism versus Realism. *Philosophy and Phenomenological Research*, 50, Supplement (Fall 1990), pp. 215-36.

