

CDD: 128.2

COMO LEVAR ESTADOS MENTAIS A SÉRIO (EPIFENÔMENOS E FINGIMENTOS)

HILAN BENSUSAN

*Departamento de Filosofia,
Universidade de Brasília,
Campus Universitário Darcy Ribeiro,
70910-900 BRASÍLIA, DF
BRASIL*

hilanb@unb.br

Resumo: Kim (1998) recomenda uma forma de fisicalismo que seria, segundo ele, a única alternativa fisicalista para evitar o epifenomenalismo. Neste trabalho, mostro que esta alternativa não é viável: o fisicalismo de Kim também não consegue lidar de maneira satisfatória com o problema da causalção mental. Considerando algumas características da simulação de estados mentais e a motivação para entendermos as propriedades mentais como separáveis das físicas, esboço uma maneira de pensar nos estados mentais que evita tanto o fisicalismo quanto o epifenomenalismo.

Palavras-chave: fisicalismo; Kim; causalção mental; epifenomenalismo.

1. Às vezes a melhor maneira de fingir que você está fazendo alguma coisa é fazê-la. Sabem disto os espões (fingem bem que trabalham, passeiam com amigos ou fazem compras precisamente trabalhando, passeando com amigos e fazendo compras) e todos aqueles que para fingirem que estão entretidos fazendo alguma coisa, aprendem a fazê-la. Austin (1957) estuda o fingimento em casos como o dos larápios que, pegos a roubar uma casa, fingem que estão cortando uma árvore ou do membro de uma quadrilha de assaltos a museus que, sob suspeita de que

está examinado o valor das peças, finge que está a limpar os vidros das vitrines do museu. Nestes casos, a melhor maneira de fingir é cortar bem as árvores e limpar bem os vidros. Austin faz notar que as vezes a existência de certos limites fazem com que o fingimento seja entendido como tal.¹ Se alguém está fingindo irritação e começa a quebrar os móveis ao seu redor há motivos para suspeitar que a irritação não é apenas fingida. O fingimento, muitas vezes, requer que haja certos limites: no teatro as peças acabam, os disfarces podem ser desmascarados – exceto o *génie malin* se a idéia de um *génie malin* for inteligível. Para fingir é preciso saber cortar árvores, limpar vidros, estar irritado, mas parece que é preciso também saber o que é fingir (em cada caso).

Quando alguém está fingindo um estado mental e ultrapassa certos limites, pode ser difícil mostrar que se trata *apenas* de um fingimento – depende, na verdade, de quem queremos convencer. Quebrar os móveis da casa de alguém não pode ser explicado por uma irritação fingida para ganhar um jogo de salão, mas pode talvez ser explicado por uma irritação fingida para convencer os vizinhos que o barulho incomoda. Os estados mentais fingidos têm um poder explicativo seguramente mais limitado do que os estados mentais eles mesmos: certas atitudes não podem ser explicadas por fingimento (“por que ela precisaria ir tão longe no seu fingimento, para convencer quem?”). Trata-se, quando fingimos um estado mental, de convencer

¹ Ron Chrisley, em conversas, sugere que o fingimento, dentro de certos limites, é a melhor maneira de lidar com o paradoxo de Newcomb. Ele pensa no paradoxo formulado assim: Dadas duas caixas, *A*, aberta com 1000 reais e outra, *B*, fechada, ou com um milhão ou com nada dentro, você pode escolher pegar *B* ou *A* e *B*. Mas antes da escolha, alguém prevê sua decisão e põe um milhão em *B* se a previsão é que você vai escolher pegar só *B* e não põe nada se a previsão é que você vai escolher pegar ambas. Acontece que as previsões são feitas por alguém que tem uma capacidade preditiva alta, induzida de experiências passadas (cf. Erickson & Fossa (1998)). Ron Chrisley pensa que se você finge que é ousado e não ganancioso, *B* terá um milhão e depois você pode escolher pegar as duas caixas.

alguém, portanto de fazer com que alguém pense que estamos em um estado mental M enquanto na verdade estamos em um estado mental, digamos, $F(M)$. Fazemos coisas que pareçam ser causadas por M , ainda que estejam sendo causadas por $F(M)$. Saber fingir um estado mental requer saber controlar as aparências (para alguém), saber como extraem-se causas mentais de efeitos em forma de atitudes ou de outros estados mentais – a irritação, por exemplo, tem como efeito a impaciência. Se as vezes a melhor maneira de fingir alguma coisa é fazê-la, as vezes a melhor maneira de fingir um estado mental é estar neste estado mental; ocorre que se eu roubei a sua mala ao sair do ônibus e sou confrontado, seria bom que eu estivesse convencido de que eu levei a mala por engano mas não posso estar no estado mental de convencimento-de-que-levei-a-mala-por-engano quando quiser, e então a segunda melhor opção pode ser F (convencimento-de-que-levei-a-mala-por-engano). Para fingir este estado mental, eu tento agir como se eu estivesse neste estado mental – tento simulá-lo. Como sabem os atores, é melhor simular o estado mental por trás de todas as expressões faciais do que cada uma das expressões faciais isoladamente. Uma vez simulado o estado mental, seus efeitos físicos se seguiriam.

2. De acordo com uma maneira epifenomenalista de entender M e $F(M)$, fingir um estado mental é fingir seus efeitos físicos; M , ou $F(M)$, não tem poder causal algum. A diferença entre M e $F(M)$, claro, aparece apenas no escopo mais limitado de $F(M)$. Uma tal posição parece contrária à intuição comum de que nossos estados mentais têm efeitos físicos. Kim (1998) sugere que o fantasma do epifenomenalismo ronda o fisicalismo. Este último está comprometido com a idéia de que todo efeito físico tem uma causa física – o fechamento causal do domínio físico. Idéias como o monismo anômalo e a superveniência das propriedades mentais em propriedades físicas têm que tomar propriedades físicas como as realizadoras da mentalidade, e, portanto, não satisfazem o *desideratum* de que as propriedades mentais tenham elas mesmas efeitos (físicos). Kim

apresenta então um dilema ao fisicalista: ou epifenomenalismo ou a tese de que propriedades mentais são de alguma forma idênticas a propriedades físicas. Qualquer coisa mais fraca que a identidade entre propriedades físicas e propriedades mentais, uma vez assumido o fisicalismo, leva ao epifenomenalismo. Kim, então, se ocupa em apresentar formas de estabelecer identidades entre propriedades físicas e propriedades mentais – ele recomenda que tratemos as propriedades mentais como propriedades de segunda ordem definidas por um papel causal ((1998), p. 98). Uma vez que o papel causal da propriedade mental é desempenhado por um conjunto de propriedades físicas, a propriedade mental é idêntica à um conjunto de propriedades físicas de segunda ordem.

Eu penso que o dilema é uma das muitas maneiras de enxergar que o fisicalismo é insustentável. Acho que a identidade entre propriedades que Kim postula também não satisfaz o *desideratum* da causalção mental. A identidade entre propriedades deve ser entendida como tendo um lastro em alguma correlação nômica entre as duas propriedades caso quisermos que a causalção mental seja mais efetiva que a velha *vis dormitiva* de Molière (Block (1990), Kim (1998, p. 52)): a propriedade mental causa aquilo que ela, por definição, causa. Kim então insiste que postula a identidade² entre propriedades que se correlacionam de modo nômico. Kim pensa que, uma vez que estamos falando abstratamente de propriedades mentais e propriedades físicas – e ele pensa que é legítimo fazê-lo e que é razoável tomá-las como independentes umas das outras – não precisamos nos preocupar com o problema de que não temos nenhuma especificação do papel causal que

² Trata-se, se não vale o primeiro dogma denunciado por Quine, de um postulado dentro de um sistema de hipóteses, não há uma questão de fato acerca de se as propriedades mentais são propriedades físicas (ou se elas são uma especificação de um papel causal). De fato, o discurso sobre propriedades e suas características muitas vezes parece indiferente à toda a poeira que Quine levantou.

corresponde a cada estado mental e portanto a identidade que ele postula não nos oferece nenhuma capacidade preditiva ou explicativa: continuamos tendo que usar os nossos termos mentais, ou as tais ditas propriedades mentais, para, por exemplo, fingir. Kim talvez pense que isto não é um problema, estamos interessados em encontrar uma maneira de conceber como as propriedades mentais podem ser causalmente efetivas. Mas eu penso que é já um problema sério pois supondo que possa haver uma especificação em termos de papel causal de uma propriedade como a dor, não vai haver nenhuma explicação (ou previsão) possível que envolva os efeitos da dor uma vez que a dor será entendida como o que quer que seja que causou aqueles efeitos: a explicação (e a previsão) será em termos (de propriedades ou realizadores) físicos. Explicações mentais, por exemplo, ficam descartadas. Kim pode ainda insistir que mesmo sem explicações mentais legítimas nós podemos falar de causas³. Kim pode dizer que a propriedade mental é apenas a propriedade de ter uma propriedade e que, portanto, é idêntica ao que instancia o papel causal; de alguma forma uma causa mental apenas especifica a causa física relevante na espécie em questão, dada a realizabilidade múltipla. Assim, dizer que “ x tem dor” explica o comportamento de x porque seu comportamento foi causado por uma ativação da fibra C que, na espécie X a que x pertence, instancia a dor. Sabemos disto porque é a fibra C que causa comportamentos como o de x na espécie X . Mas a definição de propriedades mentais em termos de papéis causais tira delas qualquer plausibilidade de poder causal. Se elas fossem definidas como sendo os instanciadores em particular de cada espécie, elas teriam a eficácia causal das propriedades físicas.

³ A discussão sobre explicação e causa é abordada no capítulo 3 do livro de Kim. O debate sobre causação mental freqüentemente desconsidera as diferenças entre causas, explicações e predições – a sobre-determinação, por exemplo, pode ser desejável para explicações e previsões, mas problemática se queremos falar de causação (ver Fuhrmann (2002)).

Talvez pareça que Kim poderia simplesmente abandonar a sua definição de propriedades mentais e assumir uma teoria da identidade indexada por espécies: dizer, por exemplo, que “dor é (na espécie e) ativação da fibra C ”. Neste caso, ainda poderíamos tentar explicar um comportamento que ocorre em diferentes espécies com base na dor, entendida como uma disjunção de propriedades (ativação da fibra C , ou ...). Neste caso, poderíamos postular que a disjunção tem poder causal, mas tenderíamos a atribuir o poder causal aos elementos da disjunção. Parece que neste caso, em que usamos os papéis causais apenas para reconhecer que propriedade física é dor, podemos dizer que as explicações que usam propriedades mentais podem sempre se mostrar incorretas ou insuficientes caso em uma espécie nem todos os comportamentos associados a uma propriedade mental estiverem presentes. Apelariamos, então, para o instanciador físico que é o único que pode explicar todos os efeitos físicos. Certamente, só poderíamos identificar o instanciador físico por meio de um critério que identifique a função ‘dor’. A função, ela mesma, é apenas um predicado ou um conceito e o que teria poder causal é a instanciação da qual a função é apenas uma possível descrição. Fala-se em causação mental porque o vocabulário mental especifica propriedades físicas. Esta participação nos critérios de especificação é importante para encontrarmos a causa dos comportamentos – por exemplo, nesta espécie o comportamento de gemer é causado pela ativação da fibra C que encontramos usando o conceito de dor. Ou seja, Kim pode dizer que a dor é apenas um descritor não-rígido através do qual, ele deve admitir, fazemos generalizações e inferências importantes na determinação de causas de algumas propriedades físicas. Mas, a esta altura, não há nada que possa balizar a diferença essencial, para Kim, entre propriedades e meros predicados – apenas um compromisso cego com o fisicalismo faz com que insistamos que os predicados físicos designam rigidamente propriedades. Note que todos estes problemas não aparecem se, em vez de identidades entre propriedades à moda de Kim, tivermos identidades

tais como as da teoria da identidade clássica (inspiradas na noção nageliana de redução interteórica). Neste caso, a propriedade mental teria poder explicativo, seria idêntica ao próprio realizador. A eficácia causal da propriedade mental seria a eficácia causal da propriedade física e poderíamos substituir qualquer predicado mental por qualquer predicado físico *inter alia* – conceitos que usamos para generalizar e fazer inferências seriam sempre designadores de propriedades⁴. Surge então um novo dilema para o fisicalismo: ou a teoria da identidade (clássica) ou o epifenomenalismo. Eu vou assumir, com Kim⁵, que este novo dilema nos joga nos braços do epifenomenalismo.

Porém a alternativa de Kim para salvar o fisicalismo do epifenomenalismo, que o leva a formular o primeiro dilema, também parece apresentar outros problemas. Quando consideramos propriedades mentais intencionais, ou seja, que possuem conteúdo, dizemos, por exemplo, que uma vez que você sabia que os donos da casa tinham especial apego ao vaso de porcelana, você não se aproximou dele ao participar da brincadeira de imitar uma hiena. Kim diria que o estado epistêmico pode ser entendido como um papel causal especificável, da mesma forma que Sellars (1963) tratou os pensamentos: seus conteúdos podem ser desprezados. Sem entrar na discussão sobre externalismo de conteúdos, penso que se tivermos que proceder assim para satisfazermos o *desideratum* da causação mental, teremos que sacrificar outros *desiderata*, em particular teremos que desistir de poder conceber nossos estados intencionais como sendo sobre o mundo – e portanto respondendo a ele. Ou seja, se os nossos estados com conteúdo forem entendidos em termos de papéis causais, o conteúdo dos nossos pensamentos serão epifenômenos e, portanto, a idéia de que temos imagens do mundo se

⁴ Esta designação seria rígida na construção que me parece mais plausível da teoria da identidade. Evidentemente uma construção assim terá que responder a todo tipo de objeção concernente a múltiplos realizadores.

⁵ Kim pensa que a teoria da identidade e a noção nageliana de redução encontra problemas que a descartam como alternativa (1998, pp. 92-97).

torna incompreensível. Uma vez, contudo, que sacrificamos a possibilidade de termos imagens do mundo, sacrificamos também uma parte do nosso *desideratum* original, ou seja, a possibilidade de que nossos pensamentos sobre o mundo possam causar nossas ações, por exemplo, quando investigamos. Parece que quando falamos de causalção mental, estamos falando da possibilidade de que o conteúdo das nossas crenças, enquanto conteúdo, possa fazer alguma diferença nas nossas ações. Uma vez que também este *desideratum* tem que ser sacrificado, ficamos com um dilema tautológico para o fisicalismo: ou o epifenomenalismo ou o epifenomenalismo.

3. Há filosofia da mente sem fisicalismo? Penso que sim. Na verdade, penso que precisamos examinar de que maneira nossas concepções da mente tornam possível o epifenomenalismo. Se concebemos os estados mentais como sendo caracterizáveis de um modo completamente independente de qualquer estado físico, estamos nos colocando em uma armadilha conceptual que pode tornar difícil abandonar o fisicalismo sem recuar para uma posição dualista igualmente indesejável (especialmente com respeito à possibilidade de causalção mental no domínio físico). Putnam (1999) critica a suposição de que a mente é um órgão (material ou pensante) e questiona a inteligibilidade da independência entre propriedades físicas e propriedades mentais (1999, pp. 84-91). A idéia de que essas propriedades são independentes é o que permite que possamos pensar em um mundo com propriedades físicas, mas sem propriedades mentais, que possamos conceber que a causalção mental seja uma ilusão e que possamos imaginar que a única maneira de conectar as propriedades é mostrar que elas são idênticas. Putnam insiste que a independência só é possível se abusarmos daquilo que podemos entender no uso do vocabulário mental. É difícil convencer alguém que uma criatura que ultrapasse todos os limites fingindo que tem estados mentais (ou seja, consegue ser aprovada no teste de Turing) não tem, na verdade, estados mentais. Talvez alguém dissesse que se a criatura fosse um *mobot* criado

por nós, poderíamos dizer que se trata, na verdade, apenas de estados físicos. Mas isto já supõe que a independência faz sentido. Penso que a dificuldade de convencer alguém que a criatura não tem estados mentais também não significa que os estados mentais são particularmente relativos a quem os observa, eles fazem parte da nossa maneira de conceber o mundo, tanto quanto os estados físicos, uns entrelaçados nos outros; não vale a pena tentar conceber-los separadamente.

Penso que é melhor conceber a mente como um conjunto de capacidades e disposições do nosso corpo. Através das atitudes dos outros, percebemos estados mentais. Podemos, é claro, nos enganar, pode ser que estejamos diante de um autômato ou de uma criatura que finge ter mente. Acho que uma excessiva sensibilidade acerca do risco de engano pode ter sido a origem da nossa concepção de mente como independente do corpo; porém, como frequentemente em filosofia, pagamos muito caro pela segurança se temos que postular que a mente pode escapar completamente da nossa possibilidade de exame – ela é concebida como independente de todas as atitudes e movimentos corporais que percebemos e se torna, portanto, imperceptível para nós. As propriedades físicas são então entendidas como uma espécie de máximo fator comum (McDowell (1982)) entre os casos que há mente e os que não há e tem que ser concebidas como independentes da mente. Talvez seja o suficiente que possamos distinguir entre M e $F(M)$ quando há limites em $F(M)$. Um estado mental pode ser fingido – e pode ser fingido por alguém que para fingir apresenta, ou simula, o estado mental. Ele pode ser fingido porque nós podemos percebê-lo no corpo dos outros mas também porque quem finge um estado mental conhece suficientemente o estado mental para poder fazer com que os outros o percebam.

Esse esboço de maneira de entender a mente também tem seus problemas a serem examinados e suas questões a serem elucidadas: trata-se de fazer filosofia da mente sem a concepção de fundo que torna o fisicalismo (e o epifenomenalismo) possível. Em particular, há um

problema relacionado à causação mental – como se relacionam estados com conteúdo e outros estados? Alguns estados têm outros estados como conteúdo, mas também alguns estados afetam estados com conteúdo, como? Sem espaço para elaborar estes problemas em detalhes concluo dizendo para aqueles que apreciam as várias faces do problema da causação mental que talvez o problema persista mesmo depois que estiver exorcizado o fantasma do epifenomenalismo.

REFERÊNCIAS BIBLIOGRÁFICAS

- AUSTIN, J.L. (1957). Pretending. *Proceedings of the Aristotelian Society*, volume XXXII, suppl.
- BLOCK, N. (1990). Can Mind Change the World? In: *Meaning and Method*, ed. G. Boolos (Cambridge, Cambridge University Press).
- ERICKSON, G.W. & FOSSA, J.A. (1998). *Dictionary of Paradox*. (Lanham, MD, University Press of America), pp. 137-139.
- FUHRMANN, A. (2002). Causal Exclusion Without Explanatory Exclusion, *Manuscrito XXV* (Special Number), 177-198.
- KIM, J. (1998). *Mind in a Physical World*. (Cambridge, Mass., MIT Press).
- McDOWELL, J. (1982). Criteria, Defeasibility and Knowledge. *Proceedings of the British Academy*, 68.
- PUTNAM, H. (1999). *The Threefold Cord: Mind, Body and World*. (Columbia University Press).
- SELLARS, W. (1963). *Science, Perception and Reality*. (Routledge & Kegan Paul), cap. I.