

O problema mente-corpo na primeira década do século XXI: visita guiada a pontos-chave da paisagem fisicista

António Zilhão
(Universidade de Lisboa - Departamento de Filosofia)
AntonioZilhao@letras.ulisboa.pt

1. Uma evolução surpreendente

Quem, como eu, tenha seguido com grande interesse o debate em torno do chamado ‘problema mente-corpo’ nos anos 80 e 90 do século passado, considerando todos os argumentos com um espírito aberto, mas apostando fortemente no lado materialista da disputa, não pode deixar de sentir-se surpreendido quando contempla o modo como esta discussão evoluiu ao longo da primeira década do século XXI.

A causa maior desta surpresa é, sem dúvida, a verificação de que o principal púlpito materialista foi, nestes últimos anos, reocupado pela teoria da identidade tipo-tipo. Como suponho que vos lembrareis, esta teoria era, até há bem pouco tempo, sistematicamente retratada, sobretudo no interior do próprio campo materialista, como uma forma ingénua, ultrapassada e pouco sofisticada de fisicismo. Outro factor indutor de surpresa é a constatação de que a teoria da identidade tipo-tipo efectuou essa reocupação vestindo a pele de um recém-chegado altamente sofisticado. Finalmente, uma terceira causa de surpresa é a descoberta de que, entre os principais proponentes deste remoçado ponto de vista teórico, se encontram filósofos que se lhe opuseram fortemente na segunda metade do século XX. Entre estes avulta o nome de Ned Block. Não creio ser necessário lembrar aqui que Block ganhou as suas credenciais teóricas no debate em Filosofia da Mente como defensor convicto do funcionalismo, o qual ele alegava ser o

único modo de combinar adequadamente a salvaguarda de intuições materialistas de base com a necessidade de preservar alguma autonomia para a Psicologia no campo teórico.

Que aconteceu? Para tentar alcançar alguma compreensão a este respeito, procurei encontrar uma resposta clara às duas perguntas seguintes:

Surgiram argumentos positivos realmente novos capazes de lançar uma luz inesperada sobre a teoria da identidade tipo-tipo, luz essa que, por sua vez, teria o condão de realçar aspectos anteriormente insuspeitados da dita teoria?

Encontraram-se realmente novas formas de responder convincentemente às objecções contra os seus amplamente conhecidos pontos fracos?

Por estranho que isso possa parecer-vos, a conclusão da investigação que empreendi, e cujo sumário aqui vos apresento, foi a de que a resposta a qualquer uma destas duas perguntas é negativa.

Pensareis, provavelmente, que estou a deixar-me levar pelo exagero retórico. Mas não é esse o caso. Para tentar convencer-vos disso mesmo, peço então a vossa indulgência e sugiro que me sigais naquilo que pode ser descrito como uma 'visita guiada pela paisagem fisicista contemporânea'.

2. De volta à casa de partida

Como prelúdio às perplexidades conceptuais com que me deparei nesta investigação, deixai-me mencionar o supracitado exemplo dos pontos de vista mais recentes de Ned Block. No dealbar deste novo milénio, Block apareceu a defender (juntamente com Stalnaker) a posição de acordo com a qual os estados sensoriais deveriam ser identificados com tipos de estados neurais. Mais especificamente, ele argumentou que os diferentes tipos de experiência consciente deveriam ser identificados com tipos de padrão de actividade nas células piramidais do cérebro, no seguimento dos trabalhos recentes de Crick e Koch em Neurofisiologia.

Como justifica Block esta sua nova perspectiva? Uma análise atenta ao seu trabalho mostra que ele o faz recorrendo a um argumento que exhibe a estrutura de uma inferência para a melhor explicação. Devo confessar que achei isto estranho. Com efeito, interroguei-me, como pode uma inferência para a melhor explicação propor-se resolver o problema *metafísico* tradicionalmente associado à identificação de propriedades mentais com um

tipo particular de propriedades físicas? Neste passo, deixai-me lembrar-vos que, a despeito da sua alegada ingenuidade, os primeiros teóricos da identidade tipo-tipo, como, por exemplo, J. J. Smart, sempre fizeram questão de sublinhar que nenhum argumento empírico poderia determinar a nossa escolha por ou contra qualquer uma das teorias filosóficas rivais que se propunham responder ao problema mente-corpo. Em coerência com este diagnóstico, as principais considerações introduzidas por Smart para sustentar o seu próprio ponto de vista no debate foram considerações de simplicidade ontológica, e não inferências para a melhor explicação.

Deixai-me contextualizar um pouco o problema, de modo a que seja possível dar-lhe uma forma mais precisa. Dentro do campo materialista, era costume pensar-se que havia um argumento, de contornos claramente anti-materialistas, ao qual não era possível escapar. Estou a falar do argumento modal que Kripke desenvolveu no seu famoso livro *Naming and Necessity*. A forma que a sabedoria materialista do século XX descobriu de responder ao desafio posto pelo argumento de Kripke foi a de negar que a adesão ao materialismo tivesse que implicar a visão estritamente reducionista da mente e do mental, típica da teoria da identidade tipo-tipo, e que constituía o alvo primeiro do seu argumento. Dizia-se desta que se tratava apenas de uma primeira tentativa de procurar uma perspectiva materialista consistente da mente e do mental e que, enquanto primeira tentativa algo rude, ela era, de facto, vulnerável às acusações de Kripke. Todavia – dizia-se também – havia surgido, entretanto, outra abordagem materialista da mente e do mental – a abordagem funcionalista – que era mais sofisticada que a teoria da identidade tipo-tipo, e que interpretava de um modo diferente desta a relação que supostamente se verificaria entre estados mentais e estados físicos. Mais em particular, o funcionalismo caracterizava-se por defender que os termos para estados mentais deveriam ser essencialmente interpretados como descrições de papéis causais ou, em alternativa, descrições de estados funcionais de ordem superior; ora, tanto num caso como noutro, os estados mentais referidos por tais termos seriam, pelo menos em princípio, instanciáveis numa variada gama de substratos físicos. E esta interpretação do problema mente-corpo, sendo ainda materialista, mantinha-se imune ao ataque de Kripke.

Com o benefício da retrospectiva, sabemos hoje que o funcionalismo tinha o seu lote próprio de problemas, bem difíceis de resolver; um deles, e não o menor deles, era, precisamente, o de descobrir como integrar o aspecto

qualitativo das experiências subjectivas conscientes na definição funcional dos estados mentais a elas associados. Como decerto sabeis, este problema veio a revelar-se pouco menos que intratável. Uma consequência desta intratabilidade foi o surgimento da ameaça de que o aspecto propriamente sentiente da nossa vida mental iria ter que ficar de fora do quadro teórico-explicativo. Esta ameaça adquiriu uma faceta particularmente vívida nas chamadas 'objecções dos zombis' propostas por Chalmers. Por meio das mesmas, pretendia-se sublinhar, de um modo intuitivamente apelativo, o facto de que o ponto de vista funcionalista não tinha quaisquer meios conceptuais à sua disposição em termos dos quais pudesse diferenciar entre nós e a nossa vida mental, por um lado, e criaturas zombiescas e a sua vida mental, por outro lado, as quais seriam, por hipótese, funcionalmente semelhantes a nós, mas destituídas de qualquer vida consciente. As teorias funcionalistas da mente e do mental eram assim acusadas de serem inúteis, tanto de um ponto de vista explicativo, uma vez que seriam cegas ao aspecto mais notório que a vida mental teria para nós, humanos, como de um ponto de vista metafísico, uma vez que se revelariam incapazes de integrar no âmbito da perspectiva materialista um importante segmento da realidade.

Não tenho quaisquer dúvidas de que era necessário e urgente fazer algo para encontrar um modo satisfatório de sair deste beco teórico. Foi, todavia, com grande surpresa que me apercebi que alguns dos investigadores de maior nomeada nesta área estavam a responder ao problema retirando simplesmente na direcção das velhas teorias da identidade tipo-tipo. Creio que a minha surpresa se deixa facilmente justificar quando, após um momento de reflexão, constatamos que, se a solução que se propõe para o problema é essa retirada, então isso significa que estamos, simplesmente, de volta ao ponto de partida inicial!

Só por si, o sentimento de circularidade associado a esta retirada é já suficientemente desagradável. Mas, mais importante do que esse ou quaisquer outros sentimentos, é o facto de que, se somos trazidos de volta ao ponto onde começámos, então temos que voltar a enfrentar os problemas que nos fizemos de lá sair. E o mais proeminente de entre eles é, precisamente, o desafio de Kripke às versões do materialismo baseadas na teoria da identidade tipo-tipo. Parece-me claro, todavia, que não se pode responder ao desafio de Kripke recorrendo a argumentos estruturados como inferências para a melhor explicação. Com efeito, o que está em jogo no argumento de Kripke é a própria possibilidade de que as teorias da

identidade tipo-tipo possam ter sentido. E é claro que uma inferência para a melhor explicação tem que permanecer sem sentido se a teoria que estamos a tentar mostrar ser a melhor forma de explicar os factos for, ela própria, destituída de sentido.

3. O primeiro obstáculo

Para que o ponto que mencionei fique completamente claro, creio que necessito de introduzir mais algum detalhe. Assim sendo, deixai-me passar em revista, com alguma minúcia, tanto o argumento de Kripke como o seu alvo materialista.

De acordo com os proponentes originais das teorias da identidade tipo-tipo nos anos 50 do século passado – Feigl, Place e Smart – a identidade tipo-tipo que supostamente se verificaria entre estados mentais e estados neurofisiológicos deveria ser concebida à imagem e semelhança das identidades teóricas já descobertas pela ciência, tais como, por exemplo, as identidades ‘água=H₂O’ ou ‘calor=energia cinética molecular média’. Deste modo, a dor, por exemplo, seria identificada teoricamente com o disparo das chamadas fibras-C no corpo do sofredor.

A substância do argumento anti-materialista de Kripke consiste na alegação de que uma tal analogia é espúria. O seu argumento é, em traços gerais, o seguinte. Apesar de as identidades teóricas descobertas pela ciência, como as mencionadas acima, serem *a posteriori*, elas têm que ser necessárias, caso sejam verdadeiras. Esta contenção deixa-se derivar, por sua vez, da contenção de que é suposto que os termos que ocupam ambos os lados destas identidades sejam, de acordo com o vocabulário de Kripke, designadores rígidos. Ora, se dois designadores rígidos designam a mesma coisa, então a asserção de identidade que os relaciona afirma a identidade dessa coisa consigo própria. E uma tal asserção de identidade é, obviamente, necessária. Portanto, se a água é, de facto, H₂O, então é-o necessariamente; e se o calor é, de facto, energia cinética molecular média, então é-o necessariamente. Assim sendo, se a analogia entre dor, calor e água é para valer, então a dor tem que ser necessariamente idêntica ao disparo das fibras-C. Todavia, alega Kripke, não faz qualquer sentido sustentar que a alegada correlação que se verifica existir entre dores e disparo das fibras-C seja de carácter necessário. Se existe realmente uma

correlação entre a experiência da dor e o disparo das fibras-C do sofredor, essa correlação é contingente, como é o caso, por exemplo, da correlação causal; mas uma tal correlação contingente não configura um caso de identidade. Portanto, ao correlacionar a dor com o disparo das fibras-C, não se alcançou qualquer identificação teórica, e não se produziu qualquer redução das propriedades mentais a propriedades físicas.

Para que não sobrem quaisquer dúvidas acerca da natureza deste argumento, creio ser necessário clarificar um pouco melhor a razão pela qual a correlação da dor com o disparo das fibras-C, caso seja verdadeira, não pode, segundo Kripke, ser necessária. Basicamente, o que Kripke alega a este respeito é que nós temos intuições poderosas de acordo com as quais as dores poderiam ter estado associadas em nós a muitos outros mecanismos neurofisiológicos em vez de terem ficado associadas ao disparo de fibras-C e que o disparo das fibras-C poderia ter estado associado em nós à experiência de estados mentais completamente diferentes do de sentir dor em vez de ter ficado associado à experiência deste estado. Estas intuições mostrar-nos-iam, então, que a correlação actual entre a experiência das dores e o disparo das fibras-C, caso seja verdadeira, é uma correlação perfeitamente contingente, originada pelo modo particular como os organismos vivos complexos evoluíram no planeta Terra. Ora, e a despeito do que parecem ser inclinações em sentido contrário, Kripke alega que tais intuições não ocorrem nos casos em que se descobrem genuínas identidades científicas.

Creio que é necessário introduzir aqui ainda mais uma clarificação. Na realidade, e contra a alegação de Kripke, não parece ser de todo óbvio que intuições deste género não se encontrem disponíveis nos casos de identificação teórica da água com H₂O ou do calor com a energia cinética molecular média. Com efeito, e para pegar só num destes dois exemplos, parece ser claramente o caso que nós temos a intuição de que, num outro mundo possível, aquilo a que, neste mundo, chamamos 'calor' poderia ter estado correlacionado com algum outro fenómeno físico que não o aumento da energia cinética média das moléculas do meio envolvente, e que algum outro estado, a que, neste mundo, damos outro nome, poderia estar, nesse outro mundo possível, correlacionado com esse aumento. Esta intuição de contingência parece ser obviamente sensata. Por que é, então, que ela não conta como evidência a favor da tese de que as identidades teóricas da

ciência seriam igualmente contingentes? Onde reside a diferença, se alguma existe, entre os dois tipos de intuição?

Deixai-me recordar-vos a resposta de Kripke a esta questão. Ela é basicamente a seguinte. No caso das identidades teóricas reais, alcançadas pela ciência, a descoberta da relação de identidade que vigora entre os referentes dos seus termos é mediada por estados cognitivos do sujeito, os quais, apesar de apenas contingentemente associados com eles, são distintos tanto do elemento designador como do elemento designado indicados no lado esquerdo dessas identidades teóricas. Assim, a descoberta, neste mundo, da identidade que se verifica existir entre o calor e a energia cinética molecular média é mediada pelas nossas sensações de calor. E a descoberta da identidade que se verifica existir entre a água e a substância química H₂O é igualmente mediada, neste mundo, pelas nossas experiências de perceber um líquido incolor, insípido e inodoro em rios, lagos e tanques, de bebê-lo, tomar banho nele, etc. A contingência que a intuição mencionada no parágrafo anterior revela não é, então, uma contingência que se encontraria associada a uma relação de identidade que vigoraria entre uma substância ou fenómeno natural e ele próprio. A contingência que ela revela é, antes, a contingência da conexão que verificamos existir entre um termo (e.g., ‘calor’) e o mediador semântico, de carácter cognitivo, por meio do qual a referência do mesmo é fixada. Acontece, porém, que, no caso da alegada identidade que se verifica existir entre a dor e o disparo das fibras-C do sofredor, tais intermediários cognitivos não se encontram disponíveis.

Ao tornar-nos conscientes desta distinção, Kripke leva-nos então a tomar consciência de que, quando dizemos, para exprimir o que parece ser uma intuição válida, que, num outro mundo possível, o calor poderia ser algo de diferente da energia cinética molecular média, esse é um modo inapropriado de nos exprimirmos. Na realidade, o modo apropriado de exprimir a intuição válida que subjaz a essa expressão deficiente seria o de dizer que, nesse mundo possível, a sensação que nós, neste mundo, estamos habituados a associar com o calor, ou com a energia cinética molecular média, e que exprimimos por meio da expressão ‘está calor’, não se encontraria, nesse outro mundo, correlacionada com o calor ou, o que é o mesmo, com a energia cinética molecular média.

Na realidade, as coisas podem ficar ainda mais confusas se supusermos que os habitantes desse outro mundo possível usariam a palavra ‘calor’ para referir o fenómeno que, nesse mundo possível, se encontraria

constantemente correlacionado com a sensação que nós, neste mundo, exprimimos por meio da expressão 'está calor'. Mas, como Kripke faz questão de insistir, o modo adequado de descrever esse estado de coisas seria o de dizer que, nesse outro mundo possível, a palavra 'calor' seria um designador rígido para um fenómeno físico que não o calor, fenómeno físico esse que, nesse mundo, se encontraria causalmente correlacionado com aquela sensação que nós, neste mundo, exprimimos por meio da expressão 'está calor', e que o designador rígido para o calor seria a palavra que, nesse outro mundo, seria usada em associação com a sensação que, nesse mundo, se encontraria causalmente correlacionada com o calor e que, neste mundo, se encontraria correlacionada com algum outro fenómeno físico que não o calor.

Resumindo o conjunto do argumento, então, o quadro geral que obtemos é o seguinte. No caso da alegada identificação entre a dor e o disparo das fibras-C temos uma intuição válida de acordo com a qual, mesmo assumindo que é um facto empírico que há uma correlação perfeita e constante entre a nossa experiência de episódios de dor e picos de actividade neural nas nossas fibras-C, a sensação a que chamamos 'dor' poderia, noutro mundo possível, ter estado correlacionada do mesmo modo perfeito e constante com outro fenómeno neurofisiológico que não o disparo das fibras-C. E temos também uma intuição válida de que o disparo das fibras-C poderia ter estado correlacionado, nesse outro mundo possível, com outra sensação que não a dor. As intuições aparentemente semelhantes a esta que temos no caso das identificações teóricas genuínas efectuadas pela ciência são, na realidade, dissemelhantes dela. A distinção crucial que se verifica entre estes dois tipos de intuição é a seguinte. As intuições de contingência associadas com as identificações teóricas válidas produzidas pela ciência são, na realidade, descrições enganadas de intuições válidas associadas ao carácter contingente do modo por meio do qual a referência dos designadores rígidos que figuram no lado esquerdo dessas expressões de identidade foi historicamente fixada. No caso das intuições acerca do carácter contingente da correlação que se verifica existir entre a dor e o disparo das fibras-C, essa contingência não é ilusória ou aparente, porque a origem da intuição não pode ser reconduzida à presença de um qualquer mediador contingente no estabelecimento da alegada identidade. Segundo Kripke, dor é o que o termo 'dor' significa. Não há qualquer mediador interno ou externo entre a palavra e o fenómeno mundano. Na realidade, não há qualquer fenómeno mundano: a

palavra designa rigidamente a própria sensação em vez de designar o que quer que seja de mundano com o que ela possa encontrar-se correlacionada (danos corporais, por exemplo). Todavia, se houvesse alguma identidade teórica que verdadeiramente relacionasse a dor com o disparo das fibras-C, uma tal identidade teria que ser necessária; por conseguinte, não poderia ter sido possível tanto poder-se conceber a dor como estando correlacionada com qualquer outro fenómeno neurofisiológico além do disparo das fibras-C como poder-se conceber o disparo das fibras-C como estando correlacionado com qualquer outro estado mental além da dor.

Ao fazer-nos seguir este percurso dialéctico, Kripke pretende então ter-nos mostrado duas coisas.

Primeiro, que a contenção, feita pelos teóricos originais da identidade tipo-tipo, de acordo com a qual a relação que se verificaria entre o calor e a energia cinética molecular média, ou entre a água e H₂O, seria uma identidade contingente, válida apenas no nosso mundo, em vez de uma identidade necessária, válida em todos os mundos possíveis, falha o alvo. Não há, de todo, identidades contingentes.

Segundo, que, dada a ausência de mediadores cognitivos contingentes no caso da descoberta da existência de uma correlação constante entre a dor e o disparo das fibras-C, não se coloca sequer a questão de ter sido feito um diagnóstico errado na avaliação original da intuição acima mencionada a respeito da natureza da conexão que se verificaria existir entre estes dois *relata*. Uma tal conexão é, de facto, contingente. E, como não existem quaisquer identidades contingentes, a correlação em causa não pode ser a da identidade.

Sendo este o caso, uma contradição emerge entre a intuição válida que temos do carácter contingente da correlação que se verifica existir entre a dor e o disparo das fibras-C e a tese da identidade teórica, a qual implica que uma tal correlação teria que ser necessária. Assim sendo, uma destas opções tem que ser deixada cair. De acordo com o ponto de vista de Kripke, a opção que devemos deixar cair tem que ser a segunda, uma vez que ninguém foi ainda capaz de esboçar ou, sequer, dar a entrever, um qualquer bom argumento que substanciasse a invalidade da nossa intuição original.

4. Ladear um obstáculo não é transpô-lo

Este é então o nó teórico que os defensores da renascida teoria da identidade tipo-tipo têm que enfrentar. Dado o modo como o mesmo se encontra atado, não é uma surpresa que as propostas por estes apresentadas para o desatar sejam propostas defendendo que a nossa intuição de que a correlação que se verifica existir entre a dor e o seu correlato neurofisiológico seria de carácter contingente é, na realidade, e contra Kripke, uma intuição inválida.

Vou chamar 'agnosticismo modal' a esta linha de ataque contra o desafio de Kripke. Os agnósticos modais defendem, basicamente, que as nossas intuições acerca do que é ou não concebível não constituem qualquer guia fiável para decidir questões de possibilidade metafísica. Deste ponto de vista, mesmo pressupondo que Kripke tem razão quando afirma que nós temos, de facto, uma intuição poderosa de acordo com a qual a dor e o disparo das fibras-C se encontram apenas contingentemente correlacionadas, a existência em nós de uma tal intuição é irrelevante para decidir questões acerca do modo como as coisas poderiam ou não ser.

Terão os agnósticos modais razão na sua alegação? Como sabê-lo? Em que fundamentos podemos basear os nossos juízos de que as intuições associadas às nossas capacidades de conceber um estado de coisas como possível não são um indicador fiável da necessidade metafísica? Deveremos, para o decidir, basear-nos em intuições acerca dessas intuições? E, nesse caso, o que devemos fazer com as intuições que apontam na direcção oposta? Esta é uma discussão que decorre num terreno muito armadilhado; como não quero desviar-me do meu objectivo, não vou agora entrar nela.

Vou por isso dar o benefício da dúvida ao agnosticismo modal e supor, sem mais discussão, que os agnósticos modais têm razão na sua alegação. A questão que quero realmente colocar é a seguinte: seguir-se-ão resultados substancialmente novos desta suposição? Na realidade, o que, no caso da avaliação da proposta de identificação das dores com o disparo das fibras-C, se segue do agnosticismo modal, enquanto tal, é apenas que é possível que esses *relata* sejam realmente idênticos, a despeito do facto de termos uma intuição bem distinta de que eles se encontram apenas contingentemente correlacionados, e que, portanto, não podem de todo ser idênticos. É claro que, dado o carácter agnóstico da perspectiva em causa, é também possível que, ainda de acordo com a mesma, as dores não sejam idênticas aos seus

correlatos neurais. Assim o que realmente se segue deste ponto de vista é, simplesmente, que não sabemos qual das possibilidades é, de facto, o caso.

Ora, esta parece-me ser uma resposta muito fraca a Kripke. É verdade que a objecção levantada por este último fica, por meio dela, bloqueada. Mas ela não fica, nem pouco mais ou menos, refutada. Além disso, mesmo aceitando esta resposta, Kripke continua a poder ter razão, mesmo que não tenha que tê-la. Na realidade, esta resposta constitui um lance puramente defensivo, cujo objectivo parece ser apenas o de evitar uma refutação poderosa da teoria da identidade tipo-tipo. Neste sentido, a única coisa que ela alcança é evitar a exclusão desta teoria do conjunto das posições viáveis no espaço lógico das respostas possíveis ao problema mente-corpo. Mesmo que pressuponhamos a sua validade, aquilo que dela obtemos é, simplesmente, que, tanto quanto somos capazes de saber, a teoria da identidade tipo-tipo mantém-se nesse conjunto, juntamente com todas as outras propostas de resolução do problema mente-corpo que já nos são familiares.

Mas a promessa associada ao ressurgimento desta teoria é a de que teríamos novas razões *positivas* para crer nela, e não apenas razões para operar uma suspensão agnóstica a respeito das objecções de Kripke contra ela. Que razões positivas poderão ser essas? Repare-se que, se, de acordo com o agnosticismo modal, não podemos recorrer às nossas intuições de concebibilidade para aceitar as objecções de Kripke, então tão-pouco deveremos poder recorrer a elas para comparar umas com as outras as diferentes sugestões de resposta ao problema metafísico que nos ocupa.

Chegados aqui, creio ser necessário enfatizar que uma das condições necessárias para a admissibilidade ao debate filosófico de uma sugestão de resposta a este problema é a de que a mesma seja compatível com a evidência empírica disponível. Portanto, todas as que continuam a participar nele o são. Por outro lado, deixai-me lembrar-vos que, se as intuições a que Kripke faz apelo são inúteis para o debate, então as intuições acerca de zombis também deveriam sê-lo. Mas, se este é o caso, é legítimo levantar-se a questão de saber por que é que é suposto que, de acordo com, por exemplo, Block, não aceitemos as primeiras e as suas consequências, enquanto que é suposto que aceitemos que as segundas mostraram a incapacidade do funcionalismo em lidar com o problema da consciência e, por conseguinte, de se estabelecer como uma perspectiva materialista viável. De facto, se as intuições acerca da possibilidade metafísica são espúrias,

poderíamos ter simplesmente desconsiderado como igualmente espúrias e, portanto, irrelevantes, as intuições acerca de zombis e mantido no activo uma parte substancial da parafernália funcionalista!

Este é precisamente o ponto do debate em que entra em cena a ideia de recorrer a argumentos construídos como inferências para a melhor explicação. Com efeito, se todas as posições que participam no debate são igualmente compatíveis com a evidência empírica, se as intuições de concebibilidade são inúteis para decidir questões de possibilidade metafísica, e se considerações acerca de simplicidade ontológica ou economia conceptual são extremamente dúbias, dado que tendem a incorrer em petições de princípio, então o único tipo de argumento que parece manter-se disponível para ser utilizado é, de facto, a inferência para a melhor explicação.

Como mencionei acima, creio que as inferências para a melhor explicação são notoriamente ineficientes para tratar deste tópico. Um dos problemas levantados pelo recurso a tais argumentos, e não o menor de entre eles, é o de que, neste contexto, não temos uma noção previamente definida daquilo que a expressão 'boa explicação' poderia querer dizer. Lembremo-nos que não devemos misturar a questão metafísica ou filosófica de determinar que espécie de relação formal se verifica entre os predicados pertencentes a dois tipos de discurso com questões empíricas relacionadas com o modo como deve alguma porção de nova evidência ser adequadamente tratada. Milhões e milhões de medicamentos são manufacturados todos os dias para evitar ou aliviar a dor. Não há dúvida que os cientistas que controlam os processos de fabrico na Bayer, ou noutros conglomerados farmacêuticos, conhecem todas as questões relevantes a respeito da história química subjacente aos efeitos que o ácido acetilsalicílico, por exemplo, tem nas nossas células nervosas. Mas não é isso que procuramos aqui. O que queremos saber é se o nosso *conceito* de dor deve ou não ser *identificado* com o conceito de disparo de fibras-C ou com o conceito de um estado fenomenológico ou com o conceito de outra coisa qualquer. Por referência a o que é que uma melhor explicação deve, nestas circunstâncias, ser considerada melhor? Creio que qualquer resposta a esta pergunta está destinada, neste contexto, a ser tão circular quanto o são os apelos à simplicidade.

Como seria de esperar, o problema mencionado no parágrafo anterior acaba por vir à tona quando consideramos algumas das tentativas recentes de usar a estratégia que lhe dá origem. Assim, no âmbito do terreno

epistémico aberto pelo agnosticismo modal, podemos encontrar propostas de uso de inferências para a melhor explicação no tratamento do problema mente-corpo que apontam para direcções precisamente opostas. Um bom exemplo do que acabou de ser dito são os argumentos propostos por Block (e Stalnaker), por um lado, e por Levine, por outro lado. De acordo com Block e Stalnaker, o uso deste tipo de inferência permite-nos vindicar a verdade da teoria da identidade tipo-tipo a respeito da mente consciente; de acordo com Levine, o que o uso deste tipo de inferência nos mostra é que há um fosso explicativo intransponível a separar o modo como os físicos tratam os *qualia* e o carácter subjectivo dos mesmos. Objectivamente, Block (e Stalnaker) e Levine não podem estar ambos certos. Deixai-me introduzir o desacordo entre eles, começando por rever os argumentos de Levine.

5. Um caminho que não leva ao destino

Se alguma resposta particular ao problema mente-corpo for vindicada por meio de um argumento tendo a estrutura de uma inferência para a melhor explicação, então tem que ser o caso que o poder explicativo fornecido por uma tal resposta seja realmente o maior. Mas como podemos determinar isso? A resposta de Levine a esta questão é, à primeira vista, um tanto ou quanto desconcertante. Com efeito, e depois de tudo o que já ficou dito acima acerca do ‘agnosticismo modal’, Levine alega que podemos determinar qual de duas ou mais opções tem maior poder explicativo recorrendo a... intuições de concebibilidade. Mas, na realidade, a despeito de ser um agnóstico a respeito da capacidade das intuições de concebibilidade nos revelarem alguma coisa de relevante acerca da possibilidade metafísica, Levine sustenta que essas intuições não são despiciendas. Segundo ele, elas desempenham um importante papel *epistemológico*.

De acordo com o ponto de vista de Levine, pode produzir-se um bom resultado explicativo no problema mente-corpo se a redução por meio da qual se tentar resolvê-lo nos fornecer ganhos explicativos análogos àqueles que o empreendimento reducionista nos forneceu nas ciências naturais. Com efeito, a descoberta científica das identidades teóricas ‘água=H₂O’ e ‘calor=energia cinética molecular média’ melhorou grandemente a nossa compreensão do comportamento macroscópico da água e dos fenómenos térmicos. Assim, alega Levine, se as identidades ‘dor=disparo das fibras-C’

ou 'consciência=actividade das células piramidais' forem para ser entendidas de um modo análogo ao das identidades teóricas da ciência, como os defensores da teoria da identidade tipo-tipo defendem que são, então tem que ser o caso que o seu estabelecimento acabe por nos fornecer um poder explicativo semelhante ao destas últimas. Ainda segundo ele, é com vista a determinar se este é ou não o caso que o apelo para intuições de concebibilidade se revela crucial.

Vejamos como. Diz-nos Levine que, se estas identidades teóricas fossem todas do mesmo género, então as intuições de concebibilidade que poderíamos extrair delas deveriam também ser análogas. Mas não são. Em particular, Levine defende que se a água é, de facto, H₂O, então, apesar de podermos conceber que poderia existir por aí alguma matéria aquosa que se viesse a revelar não ser H₂O (e, portanto, se viesse a revelar tão-pouco ser água), não podemos, de modo semelhante, conceber, de modo coerente, a possibilidade de que a substância H₂O poderia não ter sido matéria aquosa, sem introduzir mudanças de monta à química terrestre. Isto é, dentro dos constrangimentos impostos pela nossa teoria química, a identificação da água com H₂O não é, simplesmente, uma questão de pura definição, cujo objectivo seria a obtenção de alguma espécie de economia conceptual. É também um meio de iluminar explicativamente o comportamento previamente conhecido da água em diferentes situações macroscópicas e de ampliar este conhecimento. Todavia, as intuições de concebibilidade suscitadas pelas alegações de que a dor seria idêntica ao disparo das fibras-C, ou de que a consciência seria idêntica à actividade das células piramidais, não nos colocam numa situação análoga.

Na realidade, e como o argumento de Kripke o mostra, no caso das identificações psicofísicas deste género continuamos a ser capazes de conceber distintamente a possibilidade de que, por exemplo, o disparo das fibras-C poderia não ter estado associado a experiências dolorosas sem que, com isso, nos sintamos minimamente compelidos a introduzir mudanças importantes no nosso conhecimento neurofisiológico. Só que, em vez de nos mostrarem o que quer que seja de carácter metafísico, o que essas intuições de concebibilidade nos mostram é, segundo Levine, que nada há nessa identificação teórica que, por analogia com o caso da identificação da água com H₂O, contribua, mesmo que apenas minimamente, para *explicar* por que é que precisamente este género de mecanismo bioquímico está em nós associado ao género de experiências a que supostamente está associado.

A tese subjacente à interpretação de Levine destas nossas intuições é a de que uma boa explicação para um fenómeno deve tornar epistemologicamente necessária a descrição das suas principais características. Assim, o que, segundo ele, acontece quando tentamos identificar propriedades conscientes com propriedades físicas, do mesmo modo que identificamos substâncias e propriedades macrofísicas com elementos e propriedades microfísicos, é que, em cada um destes casos, obtemos intuições consideravelmente dissemelhantes a respeito de questões de necessidade epistemológica. Nas identidades teóricas genuínas, proporcionadas pela ciência, os redutores tornam epistemologicamente necessárias as principais propriedades dos materiais reduzidos; mas, ao contrário do que acontece nestas, as propriedades neurofisiológicas não tornam epistemologicamente necessárias as propriedades sentidas nos estados conscientes.

Em conclusão, Levine alega que tais identificações não são capazes de nos ajudar a explicar por que é que as propriedades conscientes são como são. E, se não são capazes de fazê-lo, então, ainda segundo Levine, não faz qualquer sentido defender que tais identificações poderiam ser alcançadas por meio de inferências para a melhor explicação.

Qual é a resposta de Block ao argumento de Levine? Juntamente com Stalnaker, ele argumenta que, em si mesmo, o propósito de uma vindicação de identidade não é providenciar uma qualquer explicação, mas, inversamente, impedir o surgimento da necessidade de procurá-la. Se é verdade que a dor é realmente idêntica ao disparo das fibras-C, então a questão de querer saber por que é que o disparo das nossas fibras-C estaria associado às nossas dores não deveria sequer ser levantada. Nessas circunstâncias, colocar esta questão seria algo de semelhante a colocar a questão de por que é que a água ocorre sempre em associação com H₂O. Obviamente, dada a identidade que se verifica existir entre a água e H₂O, colocar esta questão é o mesmo que colocar a questão por que é que a água é água, isto é, não faz sentido.

O recurso por Block e Stalnaker a este estratagema argumentativo não pode deixar de originar no participante atento neste debate o sentimento de que o que ele está a tentar obter por meio de uma tal jogada é um processo através do qual possa proibir-nos de formular uma questão difícil; mas vamos, por um momento, ignorar este sentimento.

Concentremo-nos antes na outra questão crucial suscitada por esta jogada de Block e Stalnaker e que fica por responder: se é isto o que se pretende obter com a identificação teórica que ele nos propõe, então por que é que a selecção filosófica de uma teoria da identidade tipo-tipo deveria resultar de uma inferência para a melhor *explicação*?

Com franqueza, creio que a resposta de Block e Stalnaker a esta questão em nada mais se apoia senão numa profissão de fé. Com efeito, o que eles acabam por dizer para lhe responder é que a investigação neurofisiológica futura encarregar-se-á de produzir um corpo de conhecimentos teoricamente consistente, no interior do qual a substituição da conversa acerca de conceitos fenomenalistas por conversa acerca dos seus equivalentes neurofisiológicos produzirá um ganho explicativo distintivo para a nossa compreensão dos fenómenos mentais.

Pode até ser que isto venha a ser verdade, mas como podemos sabê-lo agora? Será que o pomo da discórdia aqui não é precisamente a própria possibilidade de que esse possa alguma vez vir a ser o caso? E se a disputa filosófica/metafísica pode ser resolvida pelo simples aprofundamento da investigação empírica e pelas esperanças cognitivas nele depositadas, então para quê todo este espalhafato? O que estamos todos a fazer aqui afinal?

Deste modo, depois de terem posto em causa, de uma forma puramente negativa, a validade das intuições kripkeanas, ambas as linhas argumentativas do agnosticismo modal que apelam para o recurso a inferências para a melhor explicação terminam como um grande desapontamento.

6. Outro caminho

Mas, na realidade, os agnósticos modais dividem-se em dois tipos. Ao lado da linha de resposta inferencialista-explicativa ao desafio de Kripke, há também a linha de resposta a que chamarei da 'disposição positiva'. Ao contrário da estratégia da linha inferencialista-explicativa, a estratégia da linha da disposição positiva não consiste em recorrer a inferências para a melhor explicação. Os filósofos dispostos positivamente, de acordo com o modo como estou aqui a usar esta expressão, têm tão pouca confiança no significado modal/metafísico das nossas intuições de concebibilidade quanto os outros agnósticos modais. Mas acrescentam um elemento novo a essa

desconfiança. Este elemento novo é a aceitação da tese de acordo com a qual sensações qualitativas, tais como a dor, são realmente representadas por conceitos fenomenais próprios.

David Papineau é o mais significativo e coerente dos agnósticos modais ‘dispostos positivamente’. Tal como Kripke, ele defende que os termos para conceitos fenomenais conseguem de algum modo atingir a sua referência directamente, em vez de o fazerem através de intermediários cognitivos. Por outras palavras, Papineau sustenta, por um lado, que há uma conexão directa especial entre um termo com essas características e a sua referência e, por outro lado, que esta conexão directa consegue, de algum modo, estabelecer uma ligação semântica durável entre os dois. Até aqui, Papineau está, portanto, de acordo com o anti-materialista Kripke. Mas Papineau desenvolve também uma teoria fisicista positiva com o objectivo de explicar como é que uma tal relação misteriosa de referência pode surgir. É este aspecto do seu empreendimento teórico que me faz seleccionar o rótulo ‘disposto positivamente’ para o caracterizar.

Esta é uma contribuição surpreendente para o debate. Com efeito, quando consideramos o desenvolvimento do materialismo naturalista ao longo do século XX, aquilo que nele é mais saliente é a substituição, tanto da velha interpretação fenomenológica do sentido dos termos mentais da nossa linguagem, como da sua rival comportamental-fisicista, por uma interpretação funcional dos mesmos. A consequência dessa mudança para o problema mente-corpo foi, precisamente, a de que, enquanto que o modo como este problema afligia as mentes dos fisicistas dos anos 50 parecia intratável, o modo como ele passou a apresentar-se para os filósofos funcionalistas tornou-o susceptível de uma solução. Na realidade, a tarefa destes últimos veio a revelar-se ser a de encontrar um modo de integrar a interpretação funcional do discurso mental com o discurso neurofisiológico, em vez de ser a de tentar reduzir o sentido de termos que pareciam referir conteúdos fenomenológicos ao sentido de termos que referem eventos neurofisiológicos. Tanto a relação de sobreveniência do mental no físico, introduzida por Davidson, como a relação de realização que se verificaria entre propriedades (mentais) de ordem superior e propriedades (fisiológicas) de ordem inferior, introduzida por Putnam, conseguiram operar uma tal integração; deste modo, os eventos mentais poderiam ser identificados exemplar a exemplar com eventos de natureza física, mesmo que as propriedades mentais não pudessem ser identificadas tipo a tipo com

propriedades de natureza física. Como disse acima, estas soluções para o problema mente-corpo eram, por um lado, intrinsecamente plausíveis; e eram, por outro lado, imunes ao argumento modal de Kripke, uma vez que, precisamente, cortavam a ligação semântica indissolúvel, pressuposta por Kripke, entre os termos mentais e as nossas experiências fenomenológicas. A sua desvantagem consistia, porém, em que faziam desaparecer do mapa teórico as experiências fenomenológicas que constantemente associamos ao nosso uso desses termos. Isto é, por seu intermédio, a linguagem da Psicologia via-se reconciliada com o modo científico de ver o mundo, mas, ao mesmo tempo, ficava completa e definitivamente divorciada da fenomenologia.

Como também já mencionei acima, a saída de cena do funcionalismo, no final do século XX, parece ter sido causada por um regresso irresistível da fenomenologia. E uma das motivações que subjazem ao trabalho de Papineau em *Filosofia da Mente* parece ser, de facto, a de se confrontar com esse regresso. Papineau reconhece a força da queixa fenomenológica contra o funcionalismo e defende, em consonância, que uma compreensão adequada da semântica dos nossos termos mentais nos deve levar a reconhecer que tais termos têm *tanto* um aspecto fenomenal *como* um aspecto funcional. Assim, segundo ele, a questão que deve pôr-se ao filósofo fisicista deixa de ser a de substituir a velha e mais intuitiva interpretação fenomenológica do sentido dos termos mentais pela interpretação fisicista dos mesmos para passar a ser a de integrar a primeira na segunda.

Ao argumentar deste modo, Papineau traz-nos de volta a um período ainda mais recuado da *Filosofia da Mente* do século XX. Mais concretamente, ele traz-nos de volta à distinção, para que Wittgenstein e os filósofos do Círculo de Viena chamaram a atenção nos anos 30 e 40 desse século, entre o modo como devem entender-se usos de termos mentais na primeira pessoa do singular e o modo como devem entender-se usos de termos mentais na terceira pessoa do singular. A ideia principal que foi então avançada por estes filósofos foi a de que a semântica desses termos funcionava de modo diferente nas suas diferentes condições de uso. Para Carnap, essa diferença era uma diferença de referência: os usos na primeira pessoa eram usos fenomenológicos, que referiam as experiências privadas do sujeito, enquanto que os usos na terceira pessoa eram usos comportamental-fisicistas referindo eventos que ocorriam no mundo público. Para Wittgenstein, essa diferença era uma diferença no uso da linguagem:

nos seus usos na primeira pessoa, frases com conteúdo psicológico do género de 'Estou com dores de dentes' estariam a ser usadas expressivamente, como formas comportamentais mais sofisticadas do que o choro, o gemido ou o gesto de levar a mão à bochecha, mas não essencialmente diferentes destes; em contrapartida, nos usos na terceira pessoa, as frases com conteúdo psicológico estariam a ser usadas descritivamente, e estariam, de facto, a identificar o que mais tarde ficou a ser conhecido como papéis causais-funcionais.

Dado que Carnap se interessava principalmente pelo estudo da linguagem da ciência, ele acabou por considerar que os usos dos termos e expressões mentais na primeira pessoa eram descritivamente irrelevantes, mesmo que fossem comportamentalmente relevantes. E dado que Wittgenstein considerava os usos na primeira pessoa de tais palavras e frases como sendo apenas expressivos, acerca desses usos não se punha sequer a questão de tentar descortinar qual seria a natureza da referência desses termos nesses contextos: pura e simplesmente, não havia aí qualquer referência para descobrir. Assim, a questão de tentar encontrar uma referência comum para usos fenomenais (na primeira pessoa) e psicológicos (na terceira pessoa) de termos mentais não se levantava sequer. Por outro lado, para esses filósofos, a linguagem, científica ou outra, era um jogo normativo, público e intersubjectivo; assim sendo, a semântica deveria procurar os sentidos dos termos que dela faziam parte fora do domínio da experiência privada. Para alguém com inclinações místicas, como Wittgenstein, esta deveria permanecer virgem de todo o armadilhamento conceptual e deveria ser deixada livre e intocada para a contemplação e apreciação silenciosa e não mediada de todo o ser consciente.

Mas este não é, de todo, o caso de Papineau. Ele regressa à análise clássica da distinção primeira pessoa/terceira pessoa em termos de uma distinção entre conceitos fenomenais e conceitos de papel funcional/causal, mas, ao contrário do Carnap dos anos 30, ele defende que um termo como 'dor' tem uma referência única, comum a ambos os seus usos conceptuais, a saber, o evento físico que realiza o papel funcional da dor em nós, seja ele o disparo das fibras-C, ou o que quer que seja que a Neurofisiologia acabe por vir a descobrir a este respeito. Esta é uma variação engenhosa da tese clássica da identidade tipo-tipo, que, além de pretender vindicar o materialismo, pretende também, e em simultâneo, vindicar um papel causal-

explicativo para a chamada ‘Psicologia Popular’, ou seja, para uma teoria da acção humana construída a partir do silogismo prático aristotélico.

7. O segundo obstáculo

Já disse acima que, no contexto deste ensaio, não irei pôr em causa a validade da alegação anti-kripkeana de acordo com a qual as intuições de concebibilidade não têm as capacidades metafísicas que é costume serem-lhes atribuídas. Todavia, além da objecção de Kripke, há outra objecção que a perspectiva defendida por Papineau tem que enfrentar, a saber, a objecção da privacidade. Esta objecção pode ser expressa da seguinte forma: como é possível atribuir um valor semântico à ideia do estabelecimento de uma relação de referência privada e directa entre aquilo que Papineau considera serem os nossos termos para conceitos fenomenais e os seus referentes, qualquer que possa ser a sua natureza ontológica última? Note-se que, uma vez que recusa o materialismo, este não é, necessariamente, um problema para Kripke, pelo menos enquanto este não quiser confrontar-se com o problema da existência de outras mentes e do solipsismo.

Mas Papineau é um materialista; ele não pode, por isso, passar simplesmente por cima deste problema. Para lhe responder, ele chama então a nossa atenção para o facto de que as objecções que mencionam o problema da privacidade, tipicamente inspiradas no argumento da linguagem privada de Wittgenstein, pressupõem que a linguagem é um jogo social normativo. Todavia, e apesar de não usar exactamente estas palavras, Papineau defende que, desde que Chomsky introduziu com apreciável sucesso os conceitos de linguagem-I, faculdade da linguagem e gramática interna em Linguística, a nossa ideia de o que é uma linguagem humana deixou de ser a de que a mesma seria um jogo social e passou a ser a de que ela seria um órgão natural. Ora, a partir desta perspectiva mais recente de o que é a linguagem, segue-se que o modo como os itens mentais são mobilizados pelo órgão da linguagem e se relacionam com a realidade neurofisiológica subjacente é uma questão totalmente biológica acerca da mecânica da qual a filosofia pré-chomskyana da linguagem nada tem de relevante a dizer.

Tal como no caso do agnosticismo modal, esta resposta ao problema da privacidade, em vez de o enfrentar nos seus próprios termos, afirma apenas

uma mudança de território. E tal como o fiz no caso do agnosticismo modal, também aqui vou limitar-me a aceitar esta mudança de território sem mais discussão. Ponhamos então de lado a objecção da privacidade, aceitemos que Kripke e Wittgenstein estão ambos fora de jogo, e aceitemos também que quaisquer argumentos baseados no seu trabalho filosófico não constituem razões genuínas que impeçam a teoria da identidade tipo-tipo de ser uma posição viável no debate em torno do problema mente-corpo. Mas, mais uma vez, aquilo que me preocupa, peço desculpa por insistir, é encontrar razões *positivas* para crer na bondade da teoria da identidade tipo-tipo, e não propriamente encontrar razões para não a excluir à partida do âmbito do debate sério em torno do problema mente-corpo. O que tem então Papineau a dizer-nos de novo a este respeito?

A sua contribuição original para o debate consiste na apresentação de uma teoria positiva dos conceitos mentais, a qual permitiria conciliar a indissolubilidade da conexão semântica entre os termos fenomenais e os seus referentes, pressuposta por Kripke, com a defesa da teoria da identidade tipo-tipo. Deixai-me começar por dar-vos conta em traços largos desta teoria positiva. Trata-se, basicamente, de uma actualização naturalizada da Psicologia de Hume. Na realidade, Papineau diz-nos que os conceitos fenomenais que ele introduz na discussão do problema mente-corpo são conceitos que têm um componente indexical e que tanto podem ser usados imaginativamente como introspectivamente. No primeiro destes usos, o pensador traz à sua mente aquilo a que Hume chamou uma ‘cópia esmaecida’ da própria experiência; no segundo destes usos, o pensador amplifica a experiência em curso naquilo a que Hume chamou uma ‘cópia vívida’ de si mesma. Deste modo, em qualquer um dos usos possíveis de um conceito fenomenal, a experiência à qual o conceito se reporta é, ela própria, usada no pensamento; o ponto crucial associado a esta ideia é o de que pensar fenomenalmente numa experiência dá a sensação de se estar a ter a própria experiência. Ao contrário, quando pensamos no modo material, referimos descritivamente; as experiências não entram, elas próprias, no processo de pensamento.

Papineau sustenta, então, que seria uma falácia partir desta diferença entre o modo material de pensar e o modo fenomenal de pensar para argumentar que o primeiro deixaria algo de fora de um ponto de vista semântico. Ele baptizou esta nova falácia com a expressão ‘falácia antipatética’ e analisou-a em termos da falácia uso/menção. A ideia aqui

seria a de que, quando as pessoas argumentam que o modo material de pensar deixaria a experiência de fora, o que elas estariam realmente a fazer seria expressar a sua percepção do facto de que este modo de pensar, ao contrário do modo fenomenal de pensar, não usaria as próprias experiências no processo de pensamento. Mas, de acordo com Papineau, mesmo que este seja um fenómeno intrinsecamente interessante, ele não introduz qualquer novidade essencial na questão que nos ocupa. Sendo verdade que a maioria dos nossos conceitos se comporta como os conceitos materiais, tão-pouco deixa de ser o caso que, quando pensamos no modo fenomenal, e usamos, por conseguinte, um item para se representar a si mesmo, aquilo que fazemos não deixa de ser um acto de representação, tal como quando usamos símbolos arbitrários para produzir uma representação. O facto de um pensamento no modo fenomenal nos dar a sensação de estarmos a ter a experiência acerca da qual ele é não é então mais do que um efeito colateral de uma estratégia representacional particular do cérebro. Abreviando: do facto de haver mais do que uma forma possível de representar nada se segue a respeito do objecto da representação. Não haveria então nenhuma razão de princípio pela qual este objecto não poderia ser o mesmo em ambos os casos.

Na realidade, a história apresentada acima não é acerca de quaisquer termos presentes na nossa linguagem natural. Ela é acerca do modo como a nossa linguagem do pensamento, no sentido do termo originalmente introduzido por Fodor, consegue representar usando experiências qualitativas. Dada a posição teórica na qual se coloca, aquilo de que Papineau está aqui à procura é de uma teoria empiricamente verdadeira acerca do modo como o cérebro representa o mundo e, nele, os estados do corpo do qual é a unidade de processamento central. Tratar-se-ia então de mostrar como a verdade desta história empírica tornaria plausível a tese da identidade tipo-tipo. Neste sentido, a plausibilidade da tese metafísica que nos interessa testar é feita depender da verdade de uma hipótese empírica particular. Nestas circunstâncias, para avaliar dessa plausibilidade, necessitamos de tomar contacto com a evidência empírica que poderia sustentar o acerto da hipótese referida. Onde, então, se encontra ela? Tanto quanto eu saiba, em lado nenhum.

Como poderemos, então, avaliar a contenção de Papineau? Uma resposta possível a esta pergunta poderia ser a de se alegar que, não podendo ainda ser avaliada em termos de confirmação empírica, a hipótese

proposta poderia, todavia, ser provisoriamente avaliada em termos da sua plausibilidade à luz do conhecimento de fundo disponível. Mas, mesmo que nos mantenhamos no âmbito das intuições impressionísticas acerca de o que essa evidência poderá, eventualmente, vir a ser, dado esse conhecimento de fundo, não creio que o panorama que aí consigamos discernir favoreça as pretensões empíricas de Papineau. Com efeito, os nossos pensamentos acerca da dor, por exemplo, não são, em geral, dolorosos, nem sequer levemente (apesar de poderem ser assustadores). Neste sentido, não consigo ver como é que esta história representacional se poderia adequar ao que podemos imaginar acerca de o que poderia ser a evidência apropriada ao território teórico recortado pela abordagem de Papineau. Ou, dizendo o mesmo por outras palavras, não vejo que tenhamos quaisquer fundamentos empíricos para acreditar, mesmo que apenas provisoriamente, que esta história possa ser verdadeira. Lembremo-nos, por outro lado, que todas as outras posições avançadas no âmbito do problema mente-corpo nos apresentam igualmente histórias em termos das quais uma semântica compatível com elas poderia ser desenvolvida. Na ausência de evidência empírica a seu favor, por que razão, então, escolher esta? Mais uma vez me vejo aqui forçado a lembrar que o que me interessa obter são razões *positivas* para acreditar na teoria da identidade tipo-tipo. E o interesse que possamos ter em vindicá-la não constitui, obviamente, uma dessas razões.

Na realidade, e como, aliás, o próprio Papineau reconhece, o verdadeiro papel inferencial no seu argumento é desempenhado por um passo adicional de natureza conceptual. Este passo contém o desfecho do chamado 'argumento causal'. Em traços gerais, este argumento tem o seguinte aspecto. A primeira premissa afirma que os estados mentais, qualitativos ou não, têm efeitos físicos característicos. A segunda premissa afirma que todos os efeitos físicos são integralmente originados por causas físicas. A terceira premissa exclui a sobredeterminação como uma opção viável. A conclusão que se segue da conjunção destas três premissas é a conclusão fisicista de que *as causas mentais de efeitos físicos são, na realidade, eventos físicos*.

Obviamente, nenhuma das três premissas do argumento causal é neutra. A primeira premissa exclui, desde logo, o epifenomenismo. A segunda premissa é o sumário de uma tese particular em Filosofia da Física, a saber, a tese da completude da Física. Mais uma vez, uma tal tese necessita de ser justificada a partir de uma base independente. É óbvio que ela exclui desde logo o dualismo cartesiano. E a terceira tese introduz pressupostos

relevantes acerca da causação que também necessitam de ser objecto de uma argumentação independente.

Nenhuma destas premissas me incomoda. Não sou um epifenomenista, nem sou um dualista cartesiano. O que me incomoda neste argumento é que o mesmo sempre foi do nosso conhecimento. É um argumento a favor do materialismo. Enquanto tal, nada faz para discriminar entre as diferentes variedades de materialismo. É tão compatível com a teoria da identidade tipo-tipo como é compatível com o monismo anómalo ou com qualquer outra abordagem materialista que faça uso da ideia da identidade exemplar-exemplar. Seja como for, se nós já nos inclinamos de antemão para o materialismo, já sabemos o que é que, de acordo com as nossas convicções, deve ser o caso. O que queremos saber é *como* o que achamos que *deve* ser o caso *pode*, de facto, ser o caso. Porque, intuitivamente, parece estranho que o seja. E nada de novo surge aqui no horizonte a este respeito. Mais ainda, dado o carácter à partida não neutral das premissas do argumento, nenhum não-materialista alguma vez se deixará convencer por seu intermédio.

Assim, só por si, a história representacional de Papineau não nos leva muito longe. Ela aponta numa direcção, isto é, pode ajudar-nos a compreender, em abstracto, como é que representações pertencendo a diferentes modos de pensar – o modo fenomenal e o modo material – podem acabar por revelar-se ser representações de uma mesma coisa. Mas, dado o facto de situar-se num terreno puramente empírico, ela necessita de ser sustentada por alguma evidência que contribua, mesmo que apenas minimamente, para a sua confirmação. Na ausência de qualquer evidência empírica deste género, e na presença de o que parece ser evidência que a desconfirma, ficamos apenas com o argumento causal. Mas este, além de não permitir discriminar entre as diferentes variedades de materialismo, só consegue convencer aqueles que, desde o princípio, já se encontravam convencidos da justeza da sua conclusão.

8. Outro obstáculo ladeado mas não transposto

Na secção anterior, argumentei que, na ausência de qualquer evidência empírica que a sustente, a história cerebral-representacional de acordo com a qual o nosso cérebro usaria as próprias experiências que produz como

sinais referenciais não constitui qualquer argumento positivo a favor da teoria da identidade tipo-tipo. E argumentei, também, que, tanto quanto nos é possível ter uma ideia de o que essa evidência possa vir a ser, ela parece ir contra as pretensões de Papineau. Agora, porém, gostaria de ir um pouco mais longe. Nomeadamente, gostaria de defender que a tese de carácter mais geral de que a referência dos conceitos fenomenais, presumindo que tal coisa existe, e a referência dos conceitos psicológicos, poderia ser idêntica, para além de não ser suportada pela evidência disponível, parece também ser contraditória com, pelo menos, parte dela. Neste sentido, defendo que ela é implausível.

De um modo consistente com a sua perspectiva, Papineau defende, mais uma vez, que esta é uma questão inteiramente *a posteriori*. Também mais uma vez, não vou disputar a definição do território que ele escolheu para tratar desta questão; vou por isso tentar abordá-la a partir do ângulo empírico em que ele se coloca. Ora, quando consideramos esta contenção a partir desta perspectiva, e olhamos para a evidência que se pressupõe ser relevante, acabamos por nos aperceber que parte dela, pelo menos, aponta, sem margem para dúvidas, na direcção oposta à contenção de Papineau. Com efeito, existem resultados empíricos, como aqueles que foram obtidos por Libet nas suas experiências com o cérebro humano, ou como aqueles que foram obtidos em programas de investigação que tentam identificar as heurísticas rápidas e frugais subjacentes à implementação de uma proporção significativa dos nossos comportamentos e acções, que parecem sugerir que os mecanismos neurais que subservem um número importante das nossas respostas, decisões e intenções de agir funcionam antes de, e independentemente de, os nossos sentimentos, percepções e racionalizações conscientes.

Não estou em posição de atribuir um valor definitivo a estes resultados empíricos. Mas eles tocam um ponto sensível, na medida em que parecem sustentar uma teorização psicológica que nos permite atribuir sentido a importantes fragmentos da nossa experiência de nós mesmos. Estas considerações não são, obviamente, de cariz anti-fisicista. As propriedades fenomenais que caracterizam os nossos sentimentos, percepções e racionalizações conscientes têm efeitos físicos no mundo, nomeadamente, o que nos fazem dizer, acções que nós executamos por causa daquilo que nos fazem dizer, etc. Deste modo, se o argumento causal estiver, de facto, correcto, estas propriedades fenomenais terão que ser físicas.

Imaginemos, entretanto, que já conseguimos, de algum modo, não interessa agora qual, encontrar uma confirmação independente que sustente as premissas do argumento causal. Dada a validade lógica do argumento, isto significaria que teríamos encontrado uma confirmação independente para a sua conclusão, isto é, para a proposição que afirma a identidade entre propriedades fenomenais com eficácia causal e propriedades materiais. Se, nessas circunstâncias, levássemos em conta os resultados empíricos acima referidos, o que eles pareceriam então sugerir seria que, pelo menos num grande número de casos, as propriedades neurofisiológicas subjacentes com as quais as propriedades fenomenais se viriam a revelar ser idênticas seriam *distintas* das propriedades neurofisiológicas subjacentes com as quais as propriedades psicológicas, definidas em termos dos seus papéis causais/funcionais, se viriam a revelar ser idênticas.

Portanto, e pelo menos num número relevante de casos, os conceitos fenomenais, partindo, mais uma vez, do princípio algo duvidoso de que tal coisa existe, viriam a revelar-se ter uma referência diferente da referência dos conceitos psicológicos. Sob tais circunstâncias, não seria de todo uma surpresa que não se conseguisse encontrar evidência empírica capaz de sustentar a história psico-semântica proposta por Papineau.

A minha aposta, então, é que, se, como o defende a versão do fisicismo que ficou conhecida como teoria da identidade tipo-tipo, as propriedades fenomenais se vierem de facto a revelar ser identificáveis com propriedades neurofisiológicas, – e este é um *grande se* – acabaremos por ver-nos confrontados com uma Psicologia muito diferente daquela que temos hoje. Apesar de Papineau considerar este desfecho como teoricamente possível, ele desconsidera-o como sendo altamente improvável. Todavia, parece-me que a sua intuição de improbabilidade se encontra fortemente dependente do seu desejo de apresentar o seu próprio tipo de materialismo como sendo a única alternativa filosófica ao dualismo cartesiano que é igualmente capaz de preservar o essencial da psicologia do senso comum.

Portanto, para além de estar convencido que ele não apresentou quaisquer novos argumentos de carácter positivo que nos permitam ver a teoria da identidade tipo-tipo a uma nova luz, creio também que Papineau tão-pouco conseguiu convencer-nos que, caso a teoria da identidade tipo-tipo venha, de facto, a ser vindicada de algum modo que ainda não conhecemos nem imaginamos, então essa vindicação permitir-nos-ia, em simultâneo, salvar a Psicologia Popular como uma teoria explicativa da nossa acção

intencional. Pelo contrário, o que a evidência empírica disponível parece sugerir é que se uma tal vindicação vier, de algum modo, a ser possível, então ela acarretará, com toda a probabilidade, o abandono da Psicologia Popular como teoria explicativa da acção intencional.

9. Conclusão: estamos onde já estávamos

Espero ter tornado claro por que é que penso que os filósofos cujo trabalho passei aqui em revista não apresentaram quaisquer argumentos *positivos* realmente *novos* a favor do género particular de fisicismo de que tratei neste ensaio – a teoria da identidade tipo-tipo. E também por que é que penso que eles não responderam efectivamente às objecções que há meio século perseguem esta primeira variante do fisicismo moderno, tendo-se apenas limitado a abandonar o território onde a confrontação com elas é inevitável. Na realidade, creio que o problema da consciência se mantém tão duro de roer como sempre foi para os materialistas; e que, enquanto assim permanecer, não há grande esperança de que qualquer nova versão da teoria da identidade tipo-tipo nos possa fazer avançar para além do ponto onde já nos deixaram as variantes anteriores.

Referências Bibliográficas

- Block, N. & Stalnaker, R., 1999. Conceptual Analysis, Dualism and the Explanatory Gap. *Philosophical Review*, 108, 1-46.
- Carnap, R., 1931. Die physikalische Sprache als universal Sprache der Wissenschaft. *Erkenntnis*, 2, 219-241.
-- 1932/33. Psychologie in physikalischer Sprache. *Erkenntnis*, 3.
- Chalmers, D., 1995. Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, vol. 2, 3, 200-219.
-- 1996, *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, Oxford University Press.
- Chomsky, N., 1986, *Knowledge of Language: Its Nature, Origins and Use*, New York, Praeger.
-- 1998, *On Language*, New York, The New Press.
- Crick, F., 1994, *The Astonishing Hypothesis*, New York, Scribner.
- Crick, F. & Koch, C. 1990. Towards a Neurobiological Theory of Consciousness. *Seminars in Neuroscience*, 2, 263-275.
-- 1995. Are We Aware of Neural Activity in Primary Visual Cortex?. *Nature*, 375, 121-123.
- Davidson, D., 1980, *Essays on Actions and Events*, Oxford, Clarendon Press.
- Feigl, H., 1958. The 'Mental' and the 'Physical'. In: Feigl, H., Scriven, M. & Maxwell, G. (eds.), *The Minnesota Studies in the Philosophy of Science*, vol. 2, Minneapolis, University of Minnesota Press.
- Fodor, J., 1976, *The Language of Thought*, Brighton, The Harvester Press.
- Gigerenzer, G., 2000, *Adaptive Thinking: Rationality in the Real World*, Oxford, Oxford University Press.
- Gigerenzer, G., Todd, P., The ABC Group, 1999, *Simple Heuristics that Make Us Smart*, Oxford, Oxford University Press.
- Hume, D., 1739/40, *A Treatise of Human Nature: being an attempt to introduce the experimental Method of Reasoning into Moral Subjects*, London. (Reprinted: London: Penguin Books, 1984).
- Jackson, F., 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32, 127-136.
--1986. What Mary Didn't Know. *Journal of Philosophy*, 83, 291-295.
- Kim, J., 1998 *Mind in a Physical World*, Cambridge (MA), The MIT Press.
-- 2005 *Physicalism, or Something Near Enough*, Princeton (NJ), Princeton University Press.
- Kripke, S., 1980, *Naming and Necessity*, Cambridge (MA), Harvard University Press.

Levine, J., 1993, On Leaving Out What It's Like. In: Davies, M. & Humphreys, G. H. (eds.), *Consciousness – Psychological and Philosophical Essays*, Oxford, Blackwell, 121-136.

-- 2001, *Purple Haze*, Oxford, Oxford University Press.

Libet, B., 1993, *Neurophysiology of Consciousness: selected papers and new essays*, Boston (MA), Birkhauser.

-- 2004, *Mind Time. The temporal factor in consciousness*, Cambridge (MA), Harvard University Press.

Papineau, D., 1993. Consciousness, Physicalism and the Antipathetic Fallacy. *Australasian Journal of Philosophy*, 71, 2, 169-183.

-- 2002, *Thinking about Consciousness*, Oxford, Clarendon Press.

Place, U. T., 1956. Is Consciousness a Brain Process?. *British Journal of Psychology*, 47.

Putnam, H., 1975, Minds and Machines. In: Putnam, H., *Mind, Language and Reality*, Cambridge, Cambridge University Press.

-- 1975, The Nature of Mental States. In: *Philosophical Papers*, vol. 2., Cambridge, Cambridge University Press.

Smart, J. J., 1959. Sensations and Brain Processes. *Philosophical Review*, 68, 141-156.

Wittgenstein, L., 1958, *The Blue and Brown Books*, Oxford, Blackwell.

-- 1971, Notes for Lectures on 'Private Experience' and 'Sense Data'. In: Jones, O. R., (ed.), *The Private Language Argument*, London, MacMillan, 226-275.

-- 1984, *Philosophische Untersuchungen*. In *Werkausgabe - Band 1*, Frankfurt a.M., Suhrkamp.