

Será procedente o argumento de Kripke contra a teoria da identidade tipo-tipo?

Domingos Faria

df@domingosfaria.net

Universidade de Lisboa e LanCog

Resumo

O meu objetivo neste artigo é examinar criticamente o argumento de Kripke contra a teoria da identidade tipo-tipo. Assumindo a tese da necessidade da identidade, bem como a tese da designação rígida, Kripke sustenta que se a dor é idêntica à estimulação das fibras C, então a dor é necessariamente idêntica à estimulação das fibras C. No entanto, precisamente porque a proposição expressa pela frase "a dor não é idêntica à estimulação das fibras C" é uma possibilidade metafísica, Kripke conclui, por *modus tollens*, que não há identidade entre dor e estimulação das fibras C. Por isso, a teoria da identidade tipo-tipo não é uma solução bem-sucedida para o problema da mente-corpo. Este artigo tem duas partes. Na primeira parte, que é expositiva, apresento o argumento de Kripke contra a teoria da identidade tipo-tipo. Subsequentemente, ponho em causa o argumento de Kripke, argumentando que não é procedente, pois a nossa situação epistémica atual não nos permite determinar se é metafisicamente possível a dor não ser idêntica à estimulação das fibras C. Assim, Kripke não refuta a teoria da identidade tipo-tipo.

Palavras-chave: Filosofia da Mente, Identidade Tipo-Tipo, Kripke, Necessidade da Identidade, Designação Rígida.

Abstract

My aim in this paper is to critically assess Kripke's argument against the type-type identity theory. Assuming the thesis of the necessity of identity, as well as the thesis of rigid designation, Kripke holds that if pain is identical with C-fibre firing, then pain is necessarily identical with C-fibre firing. However, precisely because the proposition expressed by the sentence "pain is not identical with C-fibre firing" is a metaphysical possibility, Kripke concludes, by *modus tollens*, that pain and C-fibre firing are not identical. Therefore, the type-type identity theory is not a successful solution to the mind-body problem. So this paper has two parts. In the first part, which is expositive, I present Kripke's argument against the type-type identity theory. After that, I will dispute Kripke's argument, arguing that it is not sound, for our current epistemic situation does not allow us to determine whether it is a metaphysical possibility that pain is not identical with C-fibre firing. Thus, Kripke does not refute the type-type identity theory.

Key-words: Philosophy of Mind, Type-Type Identity, Kripke, Necessity of identity, Rigid Designation.

O problema mente-corpo, que é um dos tópicos fundamentais de filosofia da mente, é o problema de apurar as relações entre a mente (e os fenômenos mentais) e o corpo. Em resposta a este problema, o monismo defende haver somente um gênero de substância e, de acordo com o monismo materialista em particular, essa substância única é uma substância física. Nesse caso, como entender o problema mente-corpo à luz do materialismo? Uma hipótese que pretendo examinar detalhadamente é a da identidade entre a realidade mental e a realidade física, nomeadamente no que se refere à identidade tipo-tipo.

Os teóricos da identidade tipo-tipo, como Place ou Smart, quando afirmam que “a consciência é um processo que ocorre no cérebro” ou que “as sensações são processos cerebrais” estão a sustentar a seguinte tese: certos *tipos* de estados mentais são idênticos a certos *tipos* de estados cerebrais. Assim, por exemplo, todos os estados de dor exemplificam um tipo de estado mental, e este tipo de estado mental coincide com um tipo de estado cerebral (chamemos a tais estados de “estimulação das fibras C”). Portanto, sempre que alguém tem uma experiência de dor ocorre estimulação das fibras C no seu cérebro. Esta identidade tipo-tipo distingue-se de outras formas de identidade, como é o caso da identidade exemplar-exemplar, a qual não é uma identidade entre tipos mas sim entre particulares concretos.

De acordo com Place¹, as observações sobre a dor e a estimulação das fibras C são observações sobre o mesmo processo e não sobre dois processos independentes correlacionados. Mas para se compreender esta identidade é preciso distinguir entre o “é” de definição e o “é” de composição. Atendamos aos seguintes exemplos:

(a) “Um quadrado é um retângulo equilátero”.

(b) “O relâmpago é movimento de cargas elétricas”.

As frases (a) e (b) diferem no que diz respeito ao papel desempenhado por “é” em ambas. Pois, as frases de definição, como (a), são necessárias uma vez que são verdadeiras por definição; enquanto as frases de composição, como (b), são afirmações contingentes porque têm de ser verificadas por observação ou por metodologia científica. Além disso, enquanto em (a) o sujeito e predicado têm o mesmo significado, tal não acontece no caso das frases como (b). Em (b) o que acontece é que tanto o sujeito como o predicado são aplicáveis ao mesmo objeto, caracterizando-o adequadamente. Portanto, a afirmação da identidade tipo-tipo para estados mentais e estados cerebrais, segundo Place, deve ser entendida como uma afirmação do gênero (b).

¹ Cf. PLACE, Ullin T. – “Is Consciousness a Brain Process?”. In CHALMERS, David (org.) – *Philosophy of Mind: classical and contemporary readings*. Oxford: Oxford University Press, 2002, pp. 55-60.

De forma bastante aproximada e recorrendo à distinção fregiana entre sentido e referência, Smart² defende podermos dizer que as expressões “estrela da manhã” e “estrela da tarde” referem a mesma entidade (o planeta Vénus), mas não exprimem o mesmo sentido. Do mesmo modo, podemos sustentar que “dor” e “estimulação das fibras C” referem o mesmo tipo de estado ou evento, porém não exprimem o mesmo sentido. Assim, o mesmo fenómeno ou tipo de fenómeno pode ser representado de dois modos diferentes: num discurso que usa termos mentais e num discurso que usa termos físicos. Com isto pode-se afirmar que a dor parece uma coisa à pessoa que a sente e outra ao neurocientista que investiga o cérebro, mas esses dois discursos referem a mesma propriedade ou tipo e, assim, há identidade, embora tal identidade seja contingente, como defendem Place e Smart.

No entanto, ao contrário daquilo que os teóricos da identidade tipo-tipo alegam, Kripke³ defende que se há identidade, então a identidade é necessária. Assim, se a afirmação de identidade expressa na frase (b) é verdadeira, então é necessariamente verdadeira, mesmo que só a conheçamos com recurso à observação ou a métodos empíricos. O mesmo se passa relativamente às identidades psicofísicas, como a identidade entre a dor e a estimulação das fibras C. Mas, se é contingente que a dor seja a estimulação das fibras C (como supõem Place e Smart), então temos de admitir que não há qualquer identidade. Deste modo, a teoria da identidade tipo-tipo não seria uma boa resposta ao problema mente-corpo. O argumento de Kripke, canonicamente representado, é o seguinte:

- (1) Se “A” e “B” são designadores rígidos, e se o referente de “A” é idêntico ao referente de “B”, então o referente de “A” é necessariamente idêntico ao referente de “B”.
- (2) O termo “dor” é um designador rígido.
- (3) O termo “estimulação das fibras C” é um designador rígido.
- (4) Portanto, se a dor é idêntica à estimulação das fibras C, então a dor é necessariamente idêntica à estimulação das fibras C.
- (5) Porém, é metafisicamente possível que a dor não seja idêntica à estimulação das fibras C.
- (6) Logo, a dor não é idêntica à estimulação das fibras C.

² Cf. SMART, John J. C. – “Sensations and Brain Processes”. In CHALMERS, David (org.) – *Philosophy of Mind: classical and contemporary readings*. Oxford: Oxford University Press, 2002, pp. 60-68.

³ Cf. KRIPKE, Saul – *O Nomear e a Necessidade*. Trad. de Ricardo Santos e Teresa Filipe. Lisboa: Gradiva, 2012.

Será este argumento procedente? O argumento é válido, sendo que o passo (4) é uma consequência lógica das três primeiras premissas e a conclusão (6) segue-se por *modus tollens* dos passos (4) e (5). Mas serão as premissas verdadeiras? Para examinar criticamente se é isso o que sucede, há que examinar em pormenor as premissas (1) e (5).

Começemos com um detalhe da premissa (1): o que é um designador? Para Kripke⁴ “designador” é um termo comum para nomes e descrições que referem alguma coisa. As descrições definidas têm a forma de “o x tal que ϕx ” como “o vencedor da bola de ouro em 2012”; enquanto os nomes são entendidos como nomes próprios, tal como um nome de uma cidade ou de uma pessoa, por exemplo: “Lionel Messi”. Os designadores dividem-se em rígidos e não-rígidos⁵. Por um lado, um designador é rígido se designa o mesmo objeto em todos os mundos possíveis em que o objeto existe. Ou seja, o referente de um designador rígido é constante de mundo para mundo. Entre os designadores rígidos podemos distinguir quatro categorias:

- (A) *designadores rígidos fortes*: se o seu referente existe em todos os mundos possíveis, isto é, caso o objeto seja um existente necessário, como o número 2;
- (B) *designadores rígidos fracos*: se o seu referente não existe em todos os mundos possíveis, ou seja, caso o objeto seja um existente contingente, como Lionel Messi;
- (C) *designadores rígidos “de jure”*: são rígidos de pleno direito, como os nomes próprios e os termos para tipos naturais (por exemplo: “ouro”); e
- (D) *designadores rígidos “de facto”*: não são rígidos de pleno direito, como algumas descrições definidas, tal como “o menor número par positivo”.

Por outro lado, temos designadores não-rígidos quando sucede o designador não designar o mesmo objeto em todos os mundos possíveis. Por exemplo, “o vencedor da bola de ouro em 2012” designa no mundo atual Lionel Messi, mas noutra mundo possível designa Cristiano Ronaldo. Porém, apesar de ser verdade que o homem Messi podia não ter sido o vencedor da bola de ouro em 2012, não sucede que ele podia não ter sido Messi; pois “Messi” é um designador rígido. A ideia central de Kripke a este respeito é que nomes e descrições não são o mesmo tipo de designador. Enquanto os nomes são designadores rígidos (*de jure*), as descrições ou são designadores rígidos (*de facto*) ou são designadores não-rígidos.

⁴ Cf. KRIPKE, *ibidem*, p. 70.

⁵ Cf. KRIPKE, *ibidem*, pp. 64-65, 99.

Ao clarificar a distinção entre designadores rígidos e designadores não-rígidos usamos a terminologia dos mundos possíveis; mas o que é isso? Para Kripke⁶ os mundos possíveis são um instrumento útil para se poder pensar em situações contrafactuais, ou seja, em situações diferentes do mundo atual. Os mundos possíveis apresentam um curso alternativo dos acontecimentos, sendo uma possível história do mundo. Porém, os mundos possíveis não são descobertos, mas sim estipulados. Com isto, Kripke⁷ denuncia concepções erradas dos mundos possíveis em que estes são descritos de maneira puramente qualitativa e concebidos à luz da teoria das contrapartes de David Lewis. Para Kripke essa não é a maneira normal de pensarmos em situações contrafactuais. Mas porquê? De acordo com a teoria das contrapartes, nenhum objeto pode existir em mais do que um mundo possível. Ou seja, a contraparte de alguma coisa noutro mundo possível nunca é numericamente idêntica à própria coisa. Esta relação de “contraparte de” é caracterizada em termos de uma certa relação transmundial de semelhança. Assim, uma contraparte num mundo possível de um objeto é algo naquele mundo que é bastante semelhante a esse objeto. No entanto, para Kripke, esta é uma concepção que desvirtua a compreensão intuitiva de frases como:

(7) Messi poderia ter perdido a bola de ouro em 2012.

A ideia de Kripke é que interpretamos intuitivamente (7) como afirmando algo sobre Messi, nomeadamente, que *ele* tem uma certa propriedade modal: em circunstâncias diferentes, teria perdido a bola de ouro. Porém, a teoria das contrapartes interpreta (7) de forma incorreta, pois afirma algo acerca de *outra* pessoa, isto é, uma certa contraparte de Messi num mundo possível. Ora, isto parece contraintuitivo, pois quando afirmamos (7), supõe-se que o dito mundo possível contém este homem (Messi) e que, nesse mundo, esse homem perdeu a bola de ouro (podemos até conceber um mundo possível em que Messi nem sequer é jogador de futebol, dedicando-se antes à filosofia). Portanto, a ideia é que podemos apontar para o homem, Messi, e indagar o que *lhe* poderia ter acontecido, caso as circunstâncias fossem outras.

Tendo esclarecido o que são mundos possíveis e designadores rígidos, há que esclarecer a tese substancial da premissa (1): a necessidade da identidade. A ideia central de Kripke⁸ é que se as afirmações de identidade com designadores rígidos são verdadeiras então são necessariamente verdadeiras. Uma afirmação de identidade é qualquer afirmação segundo a

⁶ Cf. KRIPKE, *ibidem*, pp. 58-63.

⁷ Cf. KRIPKE, *ibidem*, pp. 93-96.

⁸ Cf. KRIPKE, *ibidem*, pp. 42-44, 161-162.

qual um objeto A é numericamente idêntico a um objeto B. Por exemplo, a afirmação de que Bernardo Soares é idêntico a Fernando Pessoa é uma afirmação de identidade deste género. Ora, se é verdade que no mundo atual A e B são numericamente idênticos, então A e B são necessariamente idênticos; ou por outras palavras: são idênticos em todos os mundos possíveis em que existem. Mas se podemos conceber mundos possíveis em que A não é numericamente idêntico a B (e se isso for concebível, tal distinção é metafisicamente possível) e se essa concebilidade não se revelar ilusória, então por *modus tollens* concluímos que A não é de todo idêntico a B.

Mas o que justifica esta premissa (1)? O argumento de Kripke⁹ parte de três premissas: o princípio da necessidade da autoidentidade, o princípio leibniziano da indiscernibilidade dos idênticos e a suposição de que A é idêntico a B. O argumento canonicamente representado, de acordo com Lowe¹⁰, é o seguinte:

(8) Para qualquer objeto X, necessariamente X é idêntico a X. [Necessidade da autoidentidade]

(9) Para quaisquer objetos X e Y, se X é idêntico a Y, então tudo o que for verdade acerca de X é também verdade acerca de Y. [Lei de Leibniz]

(10) A é idêntico a B. [Suposição]

(11) Necessariamente A é idêntico a A. [De (8)]

(12) É verdade de A que é necessariamente idêntico a A. [De (11)]

(13) Se A é idêntico a B, então tudo o que for verdade acerca de A é também verdade acerca de B. [De (9)]

(14) Tudo o que for verdade acerca de A é também verdade acerca de B. [*Modus ponens* de (10) e (13)]

(15) É verdade acerca de B que este é necessariamente idêntico a A. [De (12) e (14)]

(16) Necessariamente A é idêntico a B. [De (15)]

(17) Logo, se A é idêntico a B, então necessariamente A é idêntico a B. [A partir de (10) e (16)].

Para esclarecer melhor esta conclusão (17) da necessidade da identidade, apresento dois casos: um com nomes próprios e outro com fenómenos naturais.

Vejamos o caso dos nomes próprios: se é verdade que Pelé é Edson do Nascimento, então é necessariamente verdade que Pelé é Edson do Nascimento; ou seja, se no mundo atual “Pelé”

⁹ Cf. KRIPKE, *ibidem*, p. 42.

¹⁰ Cf. LOWE, Jonathan – *A Survey of Metaphysics*. Oxford: Oxford University Press, 2002, p. 85.

e “Edson do Nascimento” designam o mesmo homem, então designam esse mesmo homem em todos os mundos possíveis em que ele existe.

O argumento é o seguinte¹¹: “Pelé” e “Edson do Nascimento” são nomes e, por isso, são designadores rígidos. Assim, o homem a que nos referimos no mundo atual como “Pelé” é o mesmo homem a que nos referimos como “Pelé” em todos os mundos possíveis nos quais esse homem existe. Do mesmo modo, o homem a que nos referimos no mundo atual como “Edson do Nascimento” é o mesmo homem a que nos referimos com o nome “Edson do Nascimento” em todos os mundos possíveis nos quais esse homem existe. Ora, se no mundo atual “Pelé” e “Edson do Nascimento” referem o mesmo homem e se cada um destes nomes refere sempre o mesmo homem em todos os mundos possíveis nos quais esse homem existe, segue-se que em todos os mundos possíveis Pelé é Edson do Nascimento. Portanto, se Pelé é Edson do Nascimento, então necessariamente Pelé é Edson do Nascimento. Ou seja, não há uma história alternativa do mundo em que Pelé não seja Edson do Nascimento. O mesmo se passa com “Túlio” e “Cícero”, “Véspero” e “Fósforo”, ou “Miguel Torga” e “Adolfo Correia da Rocha”.

É claro que noutra mundo possível “Pelé” podia ser o nome de outro jogador de futebol, e Pelé podia ser conhecido por outro nome, independentemente de algum jogador de futebol ser conhecido como “Pelé”. Mas isso não seria um caso em que Pelé não seria Edson do Nascimento, pois usando esses nomes como os usamos no mundo atual estamos a designar sempre o mesmo homem relativamente a todos os mundos possíveis nos quais esse homem existe (mesmo naqueles em que outro homem é conhecido como “Pelé”). Em suma, usando os nomes “Pelé” e “Edson do Nascimento” como os usamos no mundo atual, se Pelé e Edson do Nascimento são um e o mesmo homem, então não há qualquer mundo possível em que sejam diferentes¹². A este propósito, e seguindo Kaplan¹³, talvez ajude distinguir o contexto da elocução (no qual é fixado o referente do nome) das circunstâncias de avaliação (em que é do mesmo objeto anteriormente fixado que falamos em situações contrafactuais).

É preciso deixar claro que a necessidade da identidade, expressa na premissa (1), não é uma relação entre nomes, mas sim uma relação de um objeto consigo próprio; pois ao dizer que “para todo o x e y , se $x=y$ então $y=x$ ” não estamos a falar de nomes para x e y , e esta afirmação continuaria a ser verdadeira mesmo que nunca se tivesse usado quaisquer nomes para referir fosse o que fosse. Assim, ao afirmar “Pelé é Edson do Nascimento” estou a dizer da pessoa que é Pelé e também Edson do nascimento, que ele é idêntico a si mesmo¹⁴.

¹¹ Cf. KRIPKE, *ibidem*, pp. 165-168.

¹² Cf. KRIPKE, *ibidem*, p. 167.

¹³ Cf. KAPLAN, David – “Demonstratives”. In ALMOG, Joseph, PERRY, John e WETTSTEIN, Howard (org.) – *Themes From Kaplan*. Oxford: Oxford University Press, 1989, p. 494.

¹⁴ Cf. KRIPKE, *ibidem*, pp. 170-171.

No entanto, ainda que seja uma verdade necessária que Pelé é Edson do Nascimento, não podemos saber isso *a priori*, mas somente *a posteriori*. Ou seja, não descobrimos essa identidade usando apenas o pensamento (por mais que pensássemos nunca o descobriríamos sem recorrer à experiência), foi preciso socorrer-me da experiência empírica, por exemplo, lendo jornais desportivos, para descobrir que o homem a quem chamam “Pelé” é o mesmo homem a quem chamam “Edson do nascimento”. Porém, na filosofia tradicional, como sucede em Quine (bem como em Place e Smart), defende-se que todas as verdades necessárias são conhecidas *a priori*. No entanto, considero que Kripke¹⁵ tem razão quando defende que Quine está errado. Para ver isso é preciso fazer as seguintes distinções:

(18) *Necessidade*: as coisas não podiam ter sido de outra maneira.

(19) *Contingência*: as coisas podiam ter sido de outra maneira.

(20) *A priori*: o que pode ser conhecido independentemente da experiência.

(21) *A posteriori*: o que não pode ser conhecido independentemente da experiência.

Em (18) e (19) estamos perante a modalidade metafísica (ou seja, o modo como as próprias coisas são), enquanto (20) e (21) correspondem à modalidade epistémica (isto é, ao modo como podemos *saber* que certas coisas são de facto verdadeiras ou não). Portanto, temos duas noções independentes de modalidade, o que permite acomodar, ao contrário do que Quine pensava, as noções de necessidade *a posteriori* e também de contingência *a priori*. O caso da identidade entre Pelé e Edson do Nascimento é um exemplo de uma verdade necessária que não é conhecível *a priori*, mas apenas *a posteriori*.

Mas por que razão as afirmações de identidade com designadores não-rígidos não exprimem identidade necessária? Considere-se:

(22) O vencedor do prémio de melhor jogador jovem de 1958 é o jogador com a camisola número 10 da seleção do Brasil no mundial da Suécia em 1958.

No mundo atual é verdade que o homem que ganhou o prémio de melhor jogador jovem de 1958 é o mesmo homem (Pelé) que jogou no Brasil com a camisola número 10 no mundial realizado na Suécia. Porém, poderia ter sucedido o jogador com a camisola número 10 do Brasil não ganhar o prémio de melhor jogador jovem. Podemos conceber um mundo possível em que quem ganhou aquele prémio em 1958 foi um jogador da seleção francesa. Porque os

¹⁵ Cf. KRIPKE, *ibidem*, pp. 82-84, 162-164.

designadores não-rígidos não designam o mesmo objeto em todos os mundos possíveis, não se obtém uma identidade necessária com a afirmação (22). Apesar de no mundo atual (22) referir o mesmo homem, em algumas situações contrafactuais designa outras pessoas. Portanto, (22) exprime uma verdade contingente, não sendo uma verdadeira identidade. No entanto, mais uma vez, estas não seriam situações nas quais Pelé não seria Edson do Nascimento.

Vejam agora o caso em que a tese da necessidade da identidade é aplicada a fenómenos naturais, como a água, o calor e a luz¹⁶. Os cientistas descobriram que a água é H₂O. Mas comumente identificamos a água como ‘o líquido incolor, inodoro e insípido, que enche os rios e os oceanos, que serve para matar a sede, entre outras coisas’; chamemos a estas propriedades de “propriedades qualitativas *L*”. Apesar de no mundo atual identificarmos e fixarmos a referência de “água” a partir de propriedades qualitativas *L*, podemos conceber situações contrafactuais em que as propriedades qualitativas *L* não são água, mas antes XYZ (como no caso da Terra Gémea de Putnam¹⁷). Além disso, poderíamos até não ter as propriedades qualitativas *L* e continuar a ter água (isto é, H₂O), como parece até ser o caso da água polimerizada¹⁸. Com isto vemos que o termo “água” não significa as propriedades qualitativas *L*. Portanto, fixar a referência é diferente de dar o sentido. Porém, tais mundos possíveis não nos dão situações em que a água não é H₂O ou em que a água seria XYZ, mas sim casos em que as propriedades qualitativas *L* não correspondem a água, correspondendo pelo contrário a XYZ. Contudo, em todas essas situações a água continuará a ser H₂O. Logo, se é verdade que água é H₂O, então é uma verdade necessária que água é H₂O.

É claro que pode haver a ilusão de que água poderia não ser H₂O; mas essa aparente contingência de identidade é explicada pela verdadeira contingência de que a água é a substância que tem as propriedades qualitativas *L*. Isto porque entre o termo “água” e o objeto exterior existe o intermédio da nossa experiência ou sensação e é esse intermediário que poderia ter sido diferente, não alterando em nada a identidade de água e H₂O. Assim, não podemos dizer que uma verdade necessária *a posteriori* poderia ter acabado por ser diferente (como é o caso da água é H₂O), mas podemos sustentar que uma afirmação qualitativamente correspondente, que é o intermediário cognitivo entre o termo e o seu referente, poderia ter sido falsa (como a situação em que as propriedades qualitativas *L* são água).

Do mesmo modo, se é verdade que o calor é energia cinética molecular média (ECMM), então é uma verdade necessária. É verdade que no dia-a-dia percebemos o calor pelas sensações

¹⁶ Cf. KRIPKE, *ibidem*, pp. 196-202.

¹⁷ Cf. PUTNAM, Hilary – “The Meaning of ‘Meaning’”. In: *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, volume 7 (1975), pp. 131-193.

¹⁸ Cf. KRIPKE, *ibidem*, p. 197.

características que produz ao tato e a que chamamos “sensação de calor”. Porém, “calor” não significa “sensação de calor”; pois podemos facilmente conceber situações contrafactuais em que as pessoas (ou outras criaturas similares) devido a algum problema ou a terem uma constituição biológica diferente, não teriam qualquer sensação de calor e, mesmo assim, o calor continuaria a existir no mundo. Além disso, podemos imaginar mundos possíveis em que as sensações de calor de criaturas semelhantes a nós não seriam provocadas pelo calor (isto é, pela ECMM), mas pela luz. Nestas situações contrafactuais não sucede que o calor não é ECMM ou que o calor é o mesmo que a luz; pelo contrário, seriam situações em que o calor não produz a sensação de calor ou em que a sensação de calor é produzida pela luz. No entanto, em todos esses mundos possíveis o calor continuaria a ser ECMM. E se pensarmos que esta verdade pode ser apenas contingentemente verdadeira é porque a confundimos com outra verdade: que a sensação de calor poderia ser produzida por outro fenómeno que não o calor (ou seja, ECMM).

Igualmente podemos ter a ilusão de contingência de que a luz poderia não ser um feixe de fotões. Contudo, se luz e feixe de fotões são um e o mesmo objeto no mundo atual, então não existe qualquer mundo possível em que luz e feixe de fotões são dois objetos distintos. O que poderia ser distinto é a luz ser aquilo que nos ajuda a ver ou a orientar; ou seja, o modo como fixamos o termo “luz” é contingente. Pois podemos conceber mundos possíveis em que a luz não nos ajudaria a ver ou a orientar; por exemplo, numa situação em que os seres humanos seriam cegos. Concomitantemente, podemos imaginar uma situação contrafactual em que aquilo que nos ajudaria a ver ou a orientar não seria a luz mas sim as ondas sonoras (podemos conceber uma situação em que, devido a algumas mudanças no processo evolutivo, em vez de olhos teríamos apenas uma sistema sofisticado de ecolocalização semelhante ao dos morcegos). Porém, tais mundos possíveis não são casos em que a luz não seria um feixe de fotões ou em que a luz seria som. Pelo contrário, seriam situações contrafactuais em que aquilo que nos ajudaria a ver ou a orientar não seria a luz mas o som. No entanto, em todos esses mundos possíveis a luz, se existir, continuaria a ser um feixe de fotões.

Em todos estes exemplos estamos perante casos de identidades necessárias *a posteriori*; e em todos a aparência de contingência revela-se ilusória, pois é explicada por afirmações qualitativamente correspondentes (que não são necessárias nem exprimem uma propriedade essencial do objeto), como a de Pelé ser o vencedor do prémio de melhor jogador jovem de 1958, como a de a água consistir na posse das propriedades qualitativas *L*, como a de o calor ser a sensação de calor, e como a de a luz ser aquilo que nos ajuda a ver ou a orientar. Assim, ao explicar-se esta ilusão de contingência, a identidade necessária entre Pelé e Edson do

Nascimento, entre água e H₂O, entre o calor e a ECMM, entre a luz e um feixe de fótons, entre outras, pode ser mantida.

Portanto, a aparência ou ilusão de contingência explica-se com recurso às seguintes estratégias¹⁹: pode-se apelar à noção de uma situação epistémica qualitativamente idêntica. Isto é:

(i) Apesar de a identidade ser necessária, como a identidade entre o calor e a ECMM, uma pessoa poderia estar numa situação epistemicamente igual à original, como numa situação em que o calor provoca a sensação de calor, e essa situação análoga poderia ser diferente.

Da mesma forma, pode-se explicar a aparente contingência com base na noção daquilo que seleciona a referência de um designador rígido:

(ii) A fixação de designadores rígidos, como “luz” e “feixe de fótons”, é mediada pelas nossas sensações ou experiências e essa mediação poderia ser diferente; porém, isso em nada afeta a necessidade da identidade.

No entanto, se a aparência de contingência não puder ser explicada, então pode-se aplicar o *modus tollens* e concluir que não existe de todo qualquer identidade (podendo existir outra relação qualquer entre os objetos).

Com as considerações anteriores em mente, vejamos o raciocínio de Kripke para chegar ao passo (4). Em primeiro lugar sustenta-se que “dor” é um designador rígido, ou seja, que designa rigidamente o tipo de estado mental que designa; assim, “se algo é uma dor, é essencialmente uma dor e parece absurdo supor que a dor poderia ter sido algum fenómeno diferente daquele que é”²⁰. Analogamente, supõe que “estimulação das fibras C” é um designador rígido. Deste modo, tendo em conta a tese da necessidade da identidade, pode-se concluir que se há identidade entre dor e estimulação das fibras C, então tal identidade tem de ser necessária. Com este passo (4) pode-se constatar, que ao contrário do que defendem Smart e Place, as identidades psicofísicas (bem como as identidades teóricas estabelecidas pela ciência), se verdadeiras, não são contingentes mas necessárias. Por outras palavras, afirma-se que a dor e a estimulação das fibras C são idênticas em todos os mundos possíveis em que existem. Portanto, quem defende a identidade tipo-tipo e aceita a tese da necessidade

¹⁹ Cf. KRIPKE, *ibidem*, p. 223.

²⁰ KRIPKE, *ibidem*, pp. 221-222.

da identidade terá de sustentar o passo (4) e, assim, estará comprometido com as duas seguintes ideias:

(23) Não há qualquer mundo possível em que uma estimulação das fibras C não seja uma dor.

(24) Não há qualquer mundo possível em que uma dor não seja uma estimulação das fibras C.

Mas, serão as afirmações (23) e (24) verdadeiras? Kripke²¹ defende que não; pois, de forma similar às intuições cartesianas, podemos conceber a estimulação das fibras C sem dor e podemos conceber a dor sem estimulação das fibras C. Além disso, partindo da suposição de que se alguma coisa for concebível então é metafisicamente possível, podemos afirmar que é metafisicamente possível a estimulação das fibras C não ser idêntica à dor e a dor não ser idêntica à estimulação das fibras C. O argumento da concebibilidade pode ser exposto da seguinte forma:

Por um lado, contra (24) podemos dizer que temos uma intuição de acordo com a qual a dor estaria correlacionada com outros eventos físicos que não a estimulação das fibras C. Podemos imaginar um mundo possível em que a evolução tivesse ocorrido de forma diferente e, assim, a dor decorreria do funcionamento de outras estruturas neurofisiológicas que não as fibras C. Assim, é concebível que os mecanismos neurais para produzir a sensação de dor poderiam ter sido outros. Portanto, parece metafisicamente possível uma dor não ser uma estimulação das fibras C.

Por outro lado, contra (23) temos a seguinte intuição: podemos conceber situações contrafactuais em que os seres humanos têm estimulação das fibras C, mas não sentem ou não têm qualquer dor. Esta intuição é corroborada pelo mundo Zombie de Chalmers²², segundo o qual podemos imaginar um mundo microfisicamente igual ao nosso (em que existem seres humanos como nós, com estimulação fibras C), mas em que não existe consciência nem sensações de dor.

Kripke²³ tem também uma imagem para ilustrar esta intuição: imaginemos a criação do mundo por Deus, semelhante à descrita no primeiro capítulo do *Génesis*. O que precisa Deus de fazer para que a identidade entre a luz e o feixe de fotões seja uma realidade? Para tal, basta a Deus criar a luz, ou seja, criar o feixe de fotões, como sucede (de acordo com o *Génesis*) no primeiro

²¹ Cf. KRIPKE, *ibidem*, pp. 216-220.

²² Cf. CHALMERS, David – *The Conscious Mind: in search of a fundamental theory*. Oxford: Oxford University Press, 1996, pp. 123-124.

²³ Cf. KRIPKE, *ibidem*, pp. 226-227.

dia da criação. E essa é uma situação em que haverá luz (isto é, feixe de fótons), mesmo não havendo observadores, além de Deus, para constatarem isso. Seria ilusório conceber que Deus precisa de fazer algo mais para que a luz seja o feixe de fótons. Contudo, o que constitui uma tarefa adicional para Deus é o facto de a luz (feixe de fótons) ser percebida pelos humanos de modo a ser aquilo que os ajuda a ver ou a orientar, como sucede (de acordo com o *Génesis*) no sexto dia da criação. Tal tarefa poderia ser realizada de forma diferente: em vez da luz, poderia ser o som o que ajudaria os humanos a ver ou a orientarem-se no mundo. Quanto à estimulação das fibras C, para Deus criar isso apenas precisa de criar seres com fibras C que possam ser estimuladas. Mas para que a estimulação das fibras C seja dor (ou seja sentida como dor) exige de Deus uma tarefa adicional; ou seja, Deus tem de fazer que as criaturas possam sentir a estimulação das fibras C como dor. Contudo, podemos conceber situações contrafactuais em que Deus faz criaturas para as quais a estimulação das fibras C se traduz em cócegas. Deste modo, é concebível e, portanto, metafisicamente possível, haver estimulação das fibras C sem que haja dor.

Se aceitarmos este argumento da concebibilidade, não podemos continuar a sustentar a identidade tipo-tipo; pois se a dor é o mesmo que a estimulação de fibras C, então essa identidade existe em todos os mundos possíveis. Mas, como vimos no argumento da concebibilidade, há mundos possíveis em que não existe essa identidade. Logo, a dor não é igual à estimulação das fibras C.

Para rejeitar esta conclusão há que resistir ao argumento da concebibilidade. A forma de o fazer é mostrar que a possibilidade de a dor não ser estimulação das fibras C, e inversamente, é ilusória. Ou seja, é preciso explicar a ilusão de contingência, tal como o fizemos no caso das identidades teóricas estabelecidas pela ciência. Assim, se o argumento da concebibilidade não é procedente (e se o que temos não é senão uma aparência de possibilidade), então a ilusão da contingência da identidade entre dor e estimulação das fibras C pode ser explicada com as estratégias (i) ou (ii). Mas será tal projeto viável?

Kripke²⁴ argumenta que não é possível a aparente contingência da identidade entre a dor e a estimulação das fibras C ser explicada recorrendo às estratégias (i) e (ii). Seguindo a estratégia (i) poderíamos dizer que quando se imagina que a dor pode não ser estimulação de fibras C, o que realmente se concebe é que, uma afirmação epistémica qualitativamente análoga, como a de que a sensação de dor é a estimulação das fibras C, poderia ser falsa. Porém, isto não explica a ilusão de contingência, pois se uma pessoa está numa situação epistemicamente análoga à situação original de ter uma sensação de dor, então essa pessoa tem uma dor; e não

²⁴ Cf. KRIPKE, *ibidem*, pp. 224-226.

ter qualquer sensação de dor é não ter de todo dor. Aqui não conseguimos distinguir entre sensação de dor e o fenómeno da dor tal como fizemos para a sensação de calor e o fenómeno do calor, uma vez que ter dor e sentir dor são a mesma coisa, ou seja, a dor não é algo que exista independentemente de uma *experiência de dor*. Assim, se podemos conceber que a dor é distinta da estimulação de fibras C, não podemos eliminar esta possibilidade apelando a uma situação epistémica qualitativamente idêntica, embora diferente noutros aspetos.

Além disso, se seguirmos a estratégia (ii) poderíamos talvez dizer que a fixação do designador rígido “dor” foi mediado pelas nossas sensações e seria apenas essa mediação que poderia ser diferente. Porém, esta explicação não funciona porque a sensação de dor, que seleciona o designador rígido “dor”, é uma propriedade essencial da dor; ou seja, não há sensação de dor sem dor e inversamente. Então, se a sensação de dor pode ocorrer sem a estimulação das fibras C, segue-se que a dor pode ser diferente da estimulação das fibras C. Assim, recorrendo a esta estratégia tão-pouco explicamos por que razão a contingência da identidade entre dor e estimulação das fibras C seria meramente aparente.

No caso do calor é diferente, pois o designador rígido foi selecionado em função da propriedade accidental de ser uma sensação de calor, e esta sensação de calor poderia ocorrer sem o fenómeno físico do calor. Com isto verificamos que o caso das identidades psicofísicas não pode ser interpretado de forma análoga ao das identidades teóricas da ciência. Enquanto no caso do calor e EMM há um intermediário entre o fenómeno exterior e o observador, e tal intermediário (isto é, a sensação de calor) poderia ser diferente do fenómeno do calor, no caso das identidades psicofísicas não há qualquer intermediário, pois a sensação de dor é a própria dor. Assim, parece que não se pode explicar ilusão de contingência da identidade entre dor e estimulação das fibras C; e se não é possível explicar isso, é porque se trata de uma genuína contingência.

Em suma, é possível conceber situações contrafactuais em que a dor existe sem a estimulação das fibras C e inversamente. Esta contingência não é meramente aparente, pois não a explicamos adequadamente recorrendo às estratégias (i) e (ii). Logo, não é ilusório conceber a existência de dor sem a estimulação das fibras C. Se o podemos conceber sem que estejamos a sofrer uma ilusão, então isso é metafisicamente possível. Fica portanto justificada a premissa (5). E com os passos (4) e (5) concluímos por *modus tollens* que, ao contrário do que defendem os teóricos da identidade tipo-tipo, não há qualquer identidade entre dor e estimulação das fibras C.

Mas será toda esta argumentação procedente? É possível pôr em causa a premissa (1) salientando que o argumento que justifica a tese da necessidade da identidade incorre na

falácia de petição de princípio. Por exemplo, Lowe²⁵ defende que o passo (12) do argumento não se segue do (11) e que, além disso, em (12) assume-se que qualquer afirmação de identidade verdadeira acerca de A é uma verdade necessária (que é praticamente a conclusão que se quer provar). Mas o que realmente se segue de (11) é algo mais fraco, nomeadamente que:

(12') É verdade de A que é necessariamente idêntico a si mesmo.

Isto porque o passo (11) resulta diretamente da premissa (8), ou seja, do princípio da necessidade da autoidentidade. Assim, é preciso substituir o passo (12) por (12'). E a partir deste último já não se pode concluir a tese da necessidade da identidade. No entanto, penso que mesmo que este argumento seja falacioso não se segue que a tese da necessidade da identidade é falsa. Aliás, considero que a premissa (1) é verdadeira à luz dos exemplos intuitivos que expus sobre Pelé e sobre a água, o calor e a luz.

Pode-se igualmente pôr em causa as premissas (2) e (3). O próprio Kripke²⁶ assume que “estimulação das fibras C” é um designador rígido, mas admite que essa suposição é arriscada, uma vez que nada sabemos, praticamente, sobre fibras C. Contudo, mesmo que não seja um designador rígido, pode-se supor que no contexto atual é usado como um designador rígido. E quanto à dor? Para Kripke, o termo “dor” é um designador rígido, tal como “calor” ou “luz”. Porém, filósofos como Lewis consideram que “dor” não é um designador rígido, pois o que faz o tipo *dor* ser dor é ter contingentemente certas causas e efeitos típicos. Como refere Lewis: “se o conceito de dor é o conceito de um estado que desempenha um certo papel causal, então qualquer estado que desempenhe esse papel é dor. [...] É uma questão contingente a que estado o conceito e a palavra se aplicam. Isso depende do que causa o quê”²⁷. Assim, uma dor poderia não ter sido dor e inversamente. Portanto, não é um designador rígido. Porém, mesmo que as premissas (2) e (3) sejam bastante controversas, estou inclinado a acolher as intuições kripkianas de que é falso que a dor poderia ser outro fenómeno mental diferente daquele que é. Daqui se segue que admito que a identidade entre dor e estimulação das fibras é necessariamente verdadeira caso tal identidade seja verdadeira; ou seja, aceito (4).

No entanto, não aceito a premissa (5) do argumento de Kripke. Ao recusar esta premissa não pretendo negar (5) e advogar a impossibilidade metafísica de haver dor sem estimulação das fibras C. Quero apenas defender que a nossa situação epistémica atual não nos permite

²⁵ Cf. LOWE, *ibidem*, pp. 86-87.

²⁶ Cf. KRIPKE, *ibidem*, pp. 221-222.

²⁷ LEWIS, David – “Mad Pain and Martian Pain”. *Philosophical Papers: volume I*. Oxford: Oxford University Press, 1983, p. 125.

determinar se a possibilidade afirmada em (5) se verifica ou não. Ou seja, não estamos em condições de saber se (5) é verdadeira ou falsa e, portanto, não podemos concluir em (6) que a dor não é idêntica à estimulação das fibras C. De modo a fundamentar esta minha tese acolho alguns pormenores da argumentação de Levine e Polger. Começemos por recapitular a diferença entre:

(25) O calor é a ECMM.

(26) A dor é a estimulação das fibras C.

Vimos anteriormente que a contingência concebível em (25) pode ser eliminada com recurso às estratégias (i) e (ii), enquanto isso não é possível em (26). Mas podemos encontrar, de acordo com Levine²⁸, outra diferença: em (25) temos uma identidade que é plenamente explicativa, pois o nosso conhecimento da físico-química torna inteiramente compreensível o fenómeno sem que algo importante fique por explicar. Porém, em (26) há uma lacuna explicativa porque algo importante não é explicado, nomeadamente, o carácter qualitativo da dor. Assim, a estimulação das fibras C deixa por explicar “a razão por que sentimos a dor como a sentimos”.

Havendo aqui uma lacuna explicativa entre um *quale* de dor e a estimulação das fibras C, permite-nos *pensar* que a relação entre os dois é meramente contingente, mesmo que tal relação seja de facto necessária. Portanto, se pela nossa atual compreensão das propriedades físicas não se explica o que é ter estimulação das fibras C e o seu carácter qualitativo, então parece concebível haver estimulação das fibras C sem dor e inversamente. Mas isto só é concebível devido à ininteligibilidade da relação entre dor e estimulação das fibras C, ou aos limites do nosso conhecimento. Porém, daqui não se segue que essa contingência concebível seja uma possibilidade metafísica, pois pode ser apenas ilusória. No estado atual do nosso conhecimento, parece que a verdade da premissa (5) é incognoscível para nós. Por isso penso que só se conseguíssemos desfazer esta lacuna explicativa, ou se estivéssemos numa situação epistémica ideal (como a perspetiva de Deus), poderíamos determinar se é metafisicamente possível ou impossível haver dor sem estimulação das fibras C.

Polger também partilha alguns pontos principais desta tese ao defender que “algumas afirmações de identidade podem parecer contingentes, mesmo não o sendo, porque não

²⁸ Cf. LEVINE, Joseph – “Materialism and Qualia: The Explanatory Gap”. In: *Pacific Philosophical Quarterly*, volume 64 (1983), p. 357.

conhecemos os critérios relevantes de identidade a aplicar”²⁹. Para sustentar esta conclusão, suponha-se que estamos a determinar se há identidade entre dois tipos de coisas: C_1 e C_2 . Imaginemos que C_1 e C_2 são na realidade idênticas, mas que não conhecemos as condições de identidade de C_1 ou C_2 (ou de ambas). Ora, se não sabemos como individuar C_1 ou C_2 , então, tanto quanto sabemos, as suas condições de identidade são tanto compatíveis como incompatíveis; e por isso não temos maneira de determinar se C_1 e C_2 são idênticas ou diferentes. Portanto, se estamos desinformados sobre as condições de identidade para C_1 e C_2 , então pode parecer que elas poderiam não ser idênticas; ou seja, podemos conceber epistemicamente que C_1 e C_2 poderiam ser diferentes (mesmo quando metafisicamente são idênticas). Desconhecendo-se as condições de identidade (isto é, não se sabendo como individuar C_1 e C_2) surge a ilusão de contingência.

Por exemplo, antes do desenvolvimento da química molecular, poderia ter parecido consistentemente que a água não tinha de ser H_2O ; pois, o desconhecimento das condições de identidade para a água deixava em aberto ser um elemento ou uma molécula qualquer. Mas com o desenvolvimento da química molecular esclareceram-se as condições de individuação da água. Apenas quando se conseguiu que as condições de identidade da água fossem mais precisamente especificadas é que se descobriu que a água é H_2O e que não poderia ser XYZ ou outra coisa qualquer. Isto porque “as condições de identidade da água são idênticas às condições de identidade de H_2O e incompatíveis com qualquer outra coisa. Tendo-se descoberto que as condições de identidade de água e H_2O são idênticas, a aparente contingência de «água é H_2O » é eliminada”³⁰.

Portanto, se a identidade “água é H_2O ” nos parece contingente é porque podemos ignorar os critérios relevantes de identidade a aplicar a água e a H_2O . Mas ao termos conhecimento das condições de identidade, constataremos que tal contingência é meramente ilusória. Seguindo esta estratégia, conseguimos explicar a ilusão de contingência sem recorrer às estratégias (i) e (ii) de Kripke (as quais dependem de um hiato entre aparência e realidade).

Analogamente, suponhamos que é verdade que dor é estimulação das fibras C. Se essa identidade é verdadeira, então será necessariamente verdadeira; contudo, tal identificação poderá parecer-nos ilusoriamente contingente porque podemos ignorar que a dor e a estimulação das fibras C têm as mesmas condições de identidade.

Porém, o problema é que por enquanto não sabemos se a dor e a estimulação das fibras C têm as mesmas condições de identidade. Pois quando se examina identidades psicofísicas, como a

²⁹ POLGER, Thomas – “The Illusion of Contingent Identity”. *Natural Minds*. Cambridge, MA: The MIT Press, 2004, p. 49.

³⁰ POLGER, *ibidem*.

descrita na afirmação (26), estamos numa posição muito semelhante à da química antes do desenvolvimento da teoria molecular. Ou seja, não sabemos como individuar com precisão, pelo menos, os estados cerebrais (como a estimulação das fibras C).

O único meio viável, de acordo com Polger³¹, para descobrirmos informação sobre os estados cerebrais são os estudos de imagem cerebral, como a utilização da técnica da ressonância magnética. Porém, as imagens obtidas por ressonância magnética ou por outras técnicas de análise do cérebro são, no seu melhor, uma representação incompleta do que se passa no cérebro. O que aí é representado constitui apenas uma pequena parte da atividade cerebral, não sendo representado o próprio estado cerebral. Além disso, estas técnicas dão-nos uma imagem suscetível de interpretações e também recorrem a estatísticas. Por isso, não fornecem condições precisas para saber o que são estados cerebrais e como individuá-los.

A este propósito, Merrell³² critica Polger por este apenas considerar os estudos de imagem cerebral, pois tais estudos não desempenham senão uma pequeníssima parte de uma teoria unificada sobre os estados cerebrais. Assim, é preciso considerar uma vasta conjunção de informações sobre todas as verdades acerca dos estados cerebrais, fornecidas não só pelas imagens cerebrais, mas também descobertas pelas ciências relevantes (como a neurologia, psicologia, estudos clínicos, evolução, entre outros). Porém, apesar de isto ajudar na tarefa de individuar os estados cerebrais, penso que ainda desconhecemos com detalhe as suas condições de identidade. Mas talvez no futuro venhamos a conhecê-las com precisão.

Ora, se desconhecemos as condições de identidade tanto para a estimulação das fibras C como para a dor e se, do mesmo modo, não sabemos se tais condições de identidade são incompatíveis entre si, então a possibilidade de haver ou não identidade entre elas permanece em aberto. Não sabemos se a relação entre dor e estimulação das fibras C é contingente, mas tão-pouco sabemos se tal relação é necessária.

A minha conclusão, de que não conhecemos o valor de verdade da premissa (5), pode-se apoiar neste argumento, inspirado nas ideias de Polger: se podemos conceber genuinamente (isto é, sem ilusão) que não há identidade entre X e Y, então sabemos que as condições de identidade para X e Y são incompatíveis. Substitua-se X e Y por “dor” e “estimulação das fibras C”, respetivamente. Assim, se podemos conceber genuinamente que não há identidade entre dor e estimulação das fibras C, então sabemos que as condições de identidade para dor e para a estimulação de fibras C são incompatíveis. Mas não sabemos que as condições de identidade para dor e para a estimulação de fibras C são incompatíveis. Logo, não podemos conceber

³¹ Cf. POLGER, *ibidem*, pp. 52-57.

³² Cf. MERRELL, Don – “Polger on the Illusion of Contingent Identity”. In: *International Journal of Philosophical Studies*, volume 19 (2011), p. 600.

genuinamente que não há identidade entre dor e estimulação das fibras C e, portanto, não podemos afirmar que é metafisicamente possível a dor não ser estimulação das fibras C, tratando-se de uma possibilidade que tem de ser deixada em aberto.

Portanto, a nossa situação epistémica atual não nos permite saber se (5) é verdadeira ou falsa e, por isso, o argumento de Kripke não me parece procedente: não prova que a teoria da identidade tipo-tipo é falsa. No entanto, com esta minha argumentação tão-pouco se prova que a teoria da identidade tipo-tipo é verdadeira. Polger salienta que “a teoria da identidade não está morta”³³. Contudo, parece-me que também não se pode afirmar que esteja viva. Simplesmente não se sabe se está viva ou morta. Só quando conseguirmos eliminar a lacuna explicativa e descobirmos as condições de identidade para a dor e para a estimulação das fibras C (se é que chegaremos a fazê-lo e supondo que a lacuna é eliminável) poderemos saber se a teoria da identidade tipo-tipo é verdadeira ou falsa³⁴.

³³ POLGER, *ibidem*, p. 70.

³⁴ Agradeço as correções e críticas de Matheus Silva, Ricardo Santos e Vítor Guerreiro, que me ajudaram a melhorar este artigo. A discussão deste texto com os professores Adriana Silva Graça e António Zilhão permitiu-me repensar alguns argumentos e evitar alguns erros. Através de uma Bolsa de Doutoramento (SFRH/BD/85051/2012) beneficieei, na realização deste trabalho, do apoio da Fundação para a Ciência e a Tecnologia.