

## Pode a neurobiologia nos ensinar algo sobre a consciência?\*

*Patricia Smith Churchland*

### Introdução

O sistema nervoso humano apresenta uma lista impressionante de capacidades complexas, incluindo as seguintes: perceber, aprender e lembrar, planejar, decidir, realizar ações, assim como as capacidades de estar acordado, adormecer, sonhar, prestar atenção e estar consciente. Embora a neurociência tenha progredido espetacularmente neste século, nós ainda não compreendemos satisfatoriamente como é que cada uma dessas capacidades emerge das redes de neurônios.<sup>1</sup> Nós não entendemos completamente como os seres humanos podem ser Conscientes, da mesma forma como também não entendemos como eles podem andar, correr, subir em árvores ou praticar salto com vara. E quando se toma distância de tudo isso, a consciência não se revela intrinsecamente mais misteriosa do que o controle motor. Em contraposição ao desapontamento causado pela constatação de que um completo entendimento ainda nos escapa, está o otimismo moderado, baseado principalmente na natureza do progresso já obtido. Pois a neurociência cognitiva já ultrapassou consideravelmente o que alguns filósofos céticos consideravam possível e parece provável que esse progresso vai continuar.

Ao assumir que a neurociência pode revelar os mecanismos físicos subjacentes às funções psicológicas, eu estou assumindo que é de fato o cérebro que realiza essas funções – que as capacidades da mente humana são, na verdade, capacidades do cérebro humano. Essa suposição, juntamente com a concomitante rejeição de almas, espíritos cartesianos ou “substâncias fantasmagóricas” existindo separadamente do cérebro, não é delirante. Ao contrário, é uma hipótese altamente provável, baseada em evidências atualmente disponíveis da física, química, neurociência e biologia evolucionária. Ao dizer que o fisicalismo é uma hipótese, eu pretendo enfatizar seu status como uma questão empírica. Eu não presumo que seja uma questão de análise conceitual, de *insight* a priori ou de fé religiosa, embora eu esteja ciente de que nem todos os filósofos estão de acordo comigo neste ponto.<sup>2</sup>

Além disso, eu estou convencida de que a estratégia correta para o entendimento de capacidades psicológicas é essencialmente reducionista, pela qual eu entendo, em termos gerais, que compreender os mecanismos

---

\* Título do original inglês: “Can Neurobiology Teach Us Anything About Consciousness?”. In: *The Nature of Consciousness*, N. Block, O. Flanagan & G. Güzeldere (eds.). Cambridge, Mass.: MIT Press, 1997. Traduzido por Saulo de Freitas Araujo.

neurobiológicos não é supérfluo, mas sim uma necessidade. Agora, se a ciência conseguirá ou não finalmente reduzir fenômenos psicológicos a fenômenos biológicos, é uma outra questão empírica. Adotar a estratégia reducionista significa tentar explicar os macroníveis (propriedades psicológicas) em termos de microníveis (propriedades de redes neurais).

O princípio fundamental por trás dessa estratégia de pesquisa é simples: se você quiser entender como uma coisa funciona, você precisa entender não apenas seu perfil comportamental, mas também seus componentes básicos e como eles estão organizados para constituir um sistema. Se você não tem os *designs* de engenharia disponíveis para referência, você recorre à engenharia reversa – a tática desmontar um dispositivo para ver como ele funciona.<sup>3</sup> Na medida em que eu estou tentando descobrir explicações do macro para o micro, eu sou uma reducionista. Pelo fato de muitos filósofos, que concordam comigo em relação à natureza cerebral da alma, rebelarem-se, entretanto, contra o reducionismo e considerarem-no ridículo ou até mesmo digno de pena, parece-me necessário começar explicando brevemente o que eu quero e, mais enfaticamente, o que eu *não* quero dizer com uma estratégia de pesquisa reducionista.<sup>4</sup>

Esclarecendo, em primeiro lugar, a parte “negativa”, eu posso dizer que eu *não* tenho em mente que uma estratégia de pesquisa reducionista signifique a adoção de uma *estratégia puramente bottom-up*. Até onde me é permitido dizer, ninguém em neurociência pensa que a maneira de entender o sistema nervoso é, primeiro, entender tudo a respeito das moléculas básicas, depois tudo sobre cada neurônio e cada sinapse, e continuar assim ponderadamente para ascender aos vários níveis de organização até se chegar, finalmente, ao nível mais alto – os processos psicológicos. E também não há nada na história da ciência que diga que uma estratégia de pesquisa é reducionista apenas se for puramente *bottom-up*. Essa caracterização é completamente errônea (**straw**). A pesquisa subjacente aos casos clássicos de reducionismo – explicação da termodinâmica em termos da mecânica estatística; da óptica em termos da radiação eletromagnética; da transmissão hereditária em termos do DNA – seguramente não se conformou a qualquer estratégia puramente *bottom-up*.

Na medida em que a neurociência e a psicologia estão relacionadas, meu ponto de vista é simplesmente o de que seria mais sábio realizar pesquisas simultaneamente em muitos níveis, do molecular até as redes, sistemas, áreas cerebrais e, é claro, o comportamento. Aqui, como em todas as áreas da ciência, hipóteses em vários níveis podem *se desenvolver conjuntamente*, corrigindo-se e interagindo.<sup>5</sup> Os neurocientistas seriam ingênuos de ignorar dados psicológicos, por exemplo, assim como os psicólogos seriam ingênuos de ignorar todos os dados neurobiológicos.

Em segundo lugar, por “estratégia de pesquisa reducionista” eu não quero dizer que haja algo desmerecedor, não científico ou detestável em relação a descrições ou capacidades de nível superior. Parece bem óbvio, para dar um exemplo, que certas propriedades rítmicas no sistema nervoso são propriedades de redes, resultantes dos traços individuais da membrana de diversos tipos de neurônio na rede, juntamente com o modo pelo qual o conjunto de neurônios interage. O reconhecimento de que algo é a face de Arafat, para dar um outro exemplo, emerge quase certamente dos perfis de responsividade dos neurônios na rede e da maneira como eles interagem. Nesse contexto, “emergência” é totalmente respeitável e não-fantasmagórica, significando apenas, numa primeira aproximação, “propriedade da rede”. Determinar precisamente o que é a propriedade da rede, para algum traço particular, exigirá naturalmente bastante esforço experimental. Além disso, dado que o comportamento neuronal é altamente não-linear, as propriedades de rede *nunca* são uma simples “soma das partes”. Elas são alguma função – alguma função *complicada* – das propriedades das partes. As capacidades de nível superior existem de fato e as descrições de nível superior são, portanto, necessárias para especificá-las.

Para quê o materialismo eliminativo, então? Porque a caracterização existente das capacidades de nível superior do cérebro humano, presente no que se chama, na falta de um termo melhor, de “folk psychology”, pode muito bem ser reconstruída à medida em que o tempo e a neurociência cognitiva avançam. Isto também é uma hipótese empírica, em favor da qual já existem evidências empíricas. A reconstrução já está acontecendo para categorias como “memória”, “atenção” e “raciocínio”.<sup>6</sup>

A possibilidade de uma revisão não trivial e mesmo a substituição de descrições de nível superior já existentes por categorias de nível superior ‘neurobiologicamente harmoniosas’ é o que torna o materialismo eliminativo *eliminativo*.<sup>7</sup> Por categorias ‘neurobiologicamente harmoniosas’ eu entendo aquelas que permitem explicações coerentes e integradas de todo o cérebro, passando por sistemas neurais, grandes redes, microredes e neurônios. Apenas um estúpido (**straw man**) afirmaria que não existem capacidades de nível superior, que não existem fenômenos de nível superior.<sup>8</sup> Nesse aspecto geral, meu raciocínio reflete simplesmente o seguinte fato: num sentido muito importante, nós não entendemos exatamente o que o cérebro realmente faz nesses níveis mais elevados. Assim, é mais adequado considerar mesmo nossas mais preciosas intuições sobre a função mente/cérebro como hipóteses revisáveis, ao invés de encará-las como certezas transcendentais absolutas ou certezas introspectivamente dadas. O reconhecimento da possibilidade de uma tal revisão faz uma enorme diferença na maneira pela qual nós conduzimos

experimentos psicológicos e neurobiológicos, e em como nós interpretamos os resultados.

### **Recusando a Meta Neurobiológica**

Durante as últimas décadas, um certo número de filósofos tem manifestado restrições em relação à meta da pesquisa reducionista de descobrir os mecanismos neurobiológicos de capacidades psicológicas, incluindo a capacidade de estar consciente. Consequentemente, pode ser útil considerar a base de algumas dessas restrições, com o objetivo de determinar se elas justificam o abandono da meta ou se elas deveriam enfraquecer nossas esperanças sobre o que poderia ser descoberto acerca da mente/cérebro. Eu vou considerar aqui três classes principais de objeções. Para ser breve, minhas respostas serão extremamente sucintas, sacrificando os detalhes em nome do objetivo principal.

#### *A meta é absurda (incoerente)*

Um conjunto de razões para condenar a estratégia de pesquisa reducionista pode ser assim resumido: “Eu simplesmente não posso imaginar que ver azul ou o sentimento de dor, por exemplo, poderia consistir em algum padrão de atividade de neurônios no cérebro” ou, mais grosseiramente, “eu não posso imaginar como você pode se tornar consciente a partir da carne”. Há, algumas vezes, uma boa ligação entre a premissa “é inimaginável” e a conclusão “é impossível”, mas, até onde me é permitido dizer, essa ligação é uma camada de poeira que esconde o núcleo falacioso do argumento.<sup>9</sup>

Dado que atualmente entendemos tão pouco como o cérebro humano executa quaisquer de suas capacidades, é completamente previsível que nós devamos ter dificuldades em imaginar os mecanismos neurais. Quando a comunidade científica era comparavelmente ignorante em assuntos como valência, elétrons, etc., os filósofos naturais não podiam imaginar como se poderia explicar a maleabilidade dos metais, a magnetizabilidade do ferro e a resistência à ferrugem do ouro, em termos de componentes subjacentes e de sua organização. Até o advento da biologia molecular, muitas pessoas pensaram que era inimaginável e, portanto, impossível, que um ser vivo pudesse consistir numa organização particular de moléculas “mortas”. “Eu não posso imaginar”, diziam os vitalistas, “como se pode obter *vida* a partir de *matéria* morta”.

Do ponto de vista de uma considerável ignorância, o fracasso em imaginar alguma possibilidade significa apenas isto: um fracasso de imaginação – uma capacidade psicológica entre outras. Ele não indica qualquer limitação metafísica em relação ao que nós podemos vir a compreender e não pode predizer significativamente nada sobre o futuro da pesquisa científica. Após refletir sobre a espantosa complexidade do

problema da termorregulação homeotérmica, eu acho que eu não consigo imaginar como o cérebro controla a temperatura do corpo em condições diversas. Eu suspeito, entretanto, que esse é um fato psicológico relativamente sem interesse, que reflete apenas meu estado atual de ignorância. Não é um fato metafísico interessante sobre o universo e nem mesmo um fato epistemológico sobre os limites do conhecimento científico.

Uma variação da proposta “eu não posso imaginar” se manifesta da seguinte maneira: “nós nunca poderemos saber ...”, “é impossível um dia entendermos” ou “está para sempre além da ciência mostrar que ...”. A idéia aqui é que o fato de algo ser inconcebível diz algo de decisivo sobre sua impossibilidade lógica ou empírica. Eu não quero dizer que tais propostas são sempre irrelevantes. Algumas vezes elas podem ser relevantes. Mas é que elas aparecem freqüentemente quando a ciência está nos estágios iniciais do estudo de um fenômeno.

O ponto importante aqui é que várias “certezas a priori” têm se revelado, ao longo da história, empiricamente inúteis, independentemente de quão óbvias e sinceras elas tenham sido em seu apogeu. A impossibilidade do espaço ser não-euclidiano, a impossibilidade de linhas paralelas convergirem no espaço real, a impossibilidade de se terem boas evidências de que alguns eventos são subdeterminados, de que alguém esteja agora sonhando ou de que o universo teve um começo – todas elas desapareceram quando nós passamos a ter um entendimento mais profundo de como as coisas são. Se aprendemos alguma coisa com as várias descobertas científicas contra-intuitivas, é que nossas intuições podem estar erradas. Nossas intuições sobre nós mesmos e sobre a maneira como nós trabalhamos também podem estar erradas. Não há base na teoria da evolução, na matemática ou em qualquer outro lugar para assumir que concepções pré-científicas sejam concepções cientificamente adequadas.

Uma terceira variação do tema “não, não, nunca” tira conclusões sobre como o mundo *deve realmente ser*, baseadas em *propriedades lingüísticas* de certas categorias centrais do modo atual de descrever o mundo. Permitam-me dar um exemplo já desgastado: “a categoria ‘mental’ tem um significado remoto – significa algo completamente diferente da categoria ‘físico’. Portanto, é absurdo falar do cérebro vendo ou sentindo, assim como é absurdo falar da mente tendo neurotransmissores ou conduzindo corrente”. De fato, esse absurdo categorial liquida a própria possibilidade de que a ciência pudesse descobrir que sentir dor é uma atividade de neurônios no cérebro. O epíteto “erro categorial” é algumas vezes considerado suficiente para revelar o óbvio absurdo do reducionismo.

Muito já foi falado sobre esse assunto,<sup>10</sup> e eu evitarei uma extensa discussão da filosofia da linguagem, ressaltando três breves pontos: (1) é bastante exagero supor que intuições na filosofia da linguagem possam ser

um guia confiável para o que a ciência pode e não pode descobrir sobre a natureza do universo; (2) os significados mudam à medida em que a ciência faz descobertas sobre o que são alguns macrofenômenos, em termos de sua composição e da dinâmica da estrutura subjacente; (3) é improvável que os cientistas parem suas pesquisas, quando informados de que suas hipóteses e teorias “soam engraçadas” em relação ao uso corrente. É mais provável que eles dirão assim: “as teorias podem parecer engraçadas para você, mas deixe-me ensinar a ciência que está por detrás dessas teorias e que nos leva a pensar que a teoria é verdadeira. Aí ela soará menos engraçada”. Pode-se notar que parecia engraçado para os contemporâneos de Copérnico dizer que a terra é um planeta e que se move; parecia engraçado dizer que o calor é movimento molecular, que o espaço físico é não-euclidiano ou que não há posição espacial absoluta. E assim por diante.

Que uma teoria cientificamente plausível pareça engraçada é um critério apenas de sua impopularidade, não de sua falsidade. As descobertas científicas de que um macrofenômeno é um resultado complexo da microestrutura e de sua dinâmica são tipicamente surpreendentes e soam tipicamente engraçadas à primeira vista. Obviamente, nada disso é evidência positiva de que nós podemos atingir uma redução dos fenômenos psicológicos aos fenômenos neurológicos. Significa apenas que, de qualquer forma, parecer engraçado não quer dizer nada.

*O objetivo é inconsistente com a “múltipla realizabilidade”*

O ponto central dessa objeção é que se um macrofenômeno pode ser o resultado de mais de um mecanismo (organização e dinâmica dos componentes), então ele não pode ser identificado com nenhum mecanismo, e portanto a redução do macrofenômeno ao (singular) microfenômeno subjacente é impossível. Essa objeção me parece totalmente desinteressante para a ciência. Permita-me novamente ignorar importantes detalhes e simplesmente resumir as principais respostas.

1- As explicações e, portanto, as reduções são relativas a domínios. Na biologia, por exemplo, pode ser proveitoso destacar primeiro os princípios gerais que explicam algum fenômeno visto em diferentes espécies, e então imaginar como explicar as diferenças interespecíficas, para depois, se desejável, explicar diferenças entre os indivíduos de uma mesma espécie. Assim, os princípios gerais de como corações ou estômagos funcionam são imaginados, baseados talvez em estudos de uma única espécie, deixando as particularidades para serem resolvidas depois. Corações de sapos, de macacos e de humanos funcionam essencialmente da mesma forma, mas há também diferenças significativas, além do tamanho, que exigem uma análise comparativa. Considere outros exemplos: (a) com a solução geral do problema da cópia, que surgiu a partir da descoberta da estrutura

fundamental do DNA, foi possível realizar investigações de como diferenças no DNA poderiam explicar certas diferenças no fenótipo; (b) com a solução geral do problema de como os neurônios enviam e recebem sinais, foi possível explorar detalhadamente as diferenças nos perfis de responsividade de diversas classes de neurônios.<sup>11</sup>

2- Assim que o mecanismo de algum processo biológico é descoberto, existe a possibilidade de se inventarem dispositivos que imitem aquele processo. Entretanto, a invenção da tecnologia para corações artificiais ou fígados artificiais não impede o progresso explicativo sobre fígados e corações reais; ela não nega o empreendimento redutivo. A possibilidade de que outro material hereditário diferente do DNA possa ser encontrado nas coisas de qualquer outro lugar do universo não afeta o plano básico de uma redução neste planeta. A ciência teria se tornado muito pobre se Crick e Watson tivessem abandonado seu projeto por causa da abstrata possibilidade do material hereditário marciano ou artificial. De fato, nós conhecemos o núcleo do mecanismo reprodutor *na Terra* – a saber, o DNA, e nós sabemos bastante sobre como ele realiza seu trabalho. Da mesma forma, a engenharia de neurônios artificiais e de redes neurais artificiais (RNAs) facilita e é facilitada por abordagens neurobiológicas de como funcionam os neurônios reais; os empreendimentos de engenharia não significam que a busca pelos princípios básicos do sistema nervoso esteja mal direcionada.

3- Na ciência, existem sempre questões a serem respondidas. Portanto, a compreensão do funcionamento geral de um mecanismo, tal como a descoberta do pareamento da base (**base-pairing**) do DNA, não deve ser confundida com o ideal utópico de uma redução completa – uma explicação completa. Descobertas sobre o funcionamento geral de algo levanta uma multidão de questões sobre seu funcionamento *detalhado* e, posteriormente, sobre os detalhes dos *detalhes*. Para marcar a incompletude das explicações, talvez seja melhor nós evitarmos a expressão ‘redução’ e substituí-la por ‘contato redutivo’. Assim, nós podemos dizer que o objetivo da neurociência é fazer um rico contato redutivo com a psicologia, à medida que as duas disciplinas se desenvolvam conjuntamente. Eu mesma tenho colocado em prática essa recomendação. Contudo, embora alguns filósofos mostrem interesse por ela, os cientistas consideram-na singularmente pedante. De qualquer modo, o ‘contato redutivo’ entre a biologia molecular e a macrobiologia tem se tornado cada vez mais rico, desde 1953, embora ainda restem muitas questões. O contato redutivo entre a psicologia e a neurociência também tem se tornado mais rico, especialmente na última década. No entanto, é preciso dizer que os

princípios básicos de como o cérebro funciona ainda são muito mal compreendidos.

4- Quais seriam precisamente as conseqüências programáticas do argumento da múltipla realizabilidade? Seria que a neurociência é *irrelevante* para o entendimento da natureza da mente humana? É óbvio que não. Que a neurociência *não é necessária* para o entendimento da mente humana? Não se pode negar, certamente, que ela é extremamente útil. Considere as descobertas relativas ao sono, à vigília e ao sonhar; as descobertas relativas aos cérebros comissurotomizados (**split brains**), às pessoas com lesões cerebrais focais, à neuroanatomia e à neurofisiologia do sistema visual, e assim por diante. Seria talvez que nós não devêssemos ter esperanças muito grandes? O que se quer dizer aqui, precisamente, com ‘muito grande’? A esperança de que nós descobriremos os princípios gerais de como o cérebro funciona? E por que seria esta uma esperança muito grande?

#### O cérebro causa a consciência

Recusar a meta reducionista e manter, ao mesmo tempo, o dualismo à distância é uma manobra que requer um grande cuidado. A estratégia de John Searle (Searle, 1992) é dizer que apesar do cérebro *causar* os estados conscientes, qualquer identificação entre estados conscientes e atividades cerebrais é absurda. Tradicionalmente, tem-se argumentado que o máximo que o reducionista pode esperar obter são *correlações* entre estados subjetivos e estados cerebrais, e que embora as correlações possam ser evidência de causalidade, elas não podem ser evidência de identidade. Searle tem aderido a essa objeção, dizendo que enquanto as identificações do tipo *a/b*, presentes em várias áreas da ciência, revelam a realidade por trás das aparências, no caso da consciência a realidade e a aparência são inseparáveis – não há realidade em relação à consciência, a não ser o que está presente à consciência. Não há, portanto, redução alguma a ser realizada.

Resumidamente, este é o motivo pelo qual a manobra de Searle não é convincente: ele não avalia por que os cientistas optam por identificações, quando elas acontecem. Dependendo dos dados, as identificações entre níveis distintos (**cross-level identifications**) – por exemplo, *a é b* – podem ser menos problemáticas e mais compreensíveis cientificamente do que supor que *a causa b*, sendo ambos distintos. Isso pode ser melhor compreendido através de um exemplo.<sup>12</sup>

A ciência, como nós a conhecemos, não diz que a corrente elétrica em um fio é causada por elétrons em movimento, mas sim que ela é elétrons em movimento. Os genes não são causados por agrupamentos (**chunks**) de pares de base (**base-pairs**) no DNA; eles *são* agrupamentos de

pares de base (embora os agrupamentos estejam, algumas vezes, distribuídos). A temperatura não é causada pela energia cinética molecular; ela *é* energia cinética molecular. Reflita por um momento sobre a criatividade necessária para gerar explicações que mantenham a *não-identidade* e a dependência causal de (a) corrente elétrica e elétrons em movimento, (b) genes e agrupamentos de DNA, e (c) calor e movimento molecular. Alguém desfamiliarizado com os relevantes dados convergentes e com os sucessos explicativos pode supor que isso não seja tão difícil. Aqui entra Betty Crocker na história.

Em seu livro de receitas para forno microondas, Betty Crocker se oferece para explicar como funciona um forno microondas. Ela diz que quando você liga o forno, as microondas agitam as moléculas de água na comida, fazendo com que elas se movam cada vez mais rápido. Mas ela termina sua explicação aqui – como qualquer professor de ciências do segundo grau sabe que ela deveria fazer – dizendo talvez que “o aumento da temperatura *é* apenas o aumento da energia cinética das moléculas constituintes”? Não, ela não termina. Ela continua a explicar que por causa das moléculas moverem-se mais rápido, elas entram mais frequentemente em colisão umas com as outras, o que aumenta a fricção entre as moléculas e, como nós sabemos, a fricção causa calor. Betty Crocker ainda pensa que o calor *é* algo diferente da energia cinética molecular; algo causado por, mas de fato independente do movimento molecular.<sup>13</sup> Por que os cientistas também não pensam assim?

Respondendo de forma direta, porque as explicações para os fenômenos do calor – produção por combustão, pelo sol e em reações químicas; da condutividade, incluindo a condutividade em um vácuo, a variação da condutividade em materiais distintos, etc. – são *muito mais* simples e mais coerentes, quando se assume que calor *é* energia molecular das moléculas constituintes. Agora, tentar adequar os dados à suposição de que o calor *é* alguma outra coisa *causada pela* aceleração do movimento molecular, *é* como tentar pregar geléia na parede.

Se você está disposto a aderir à termodinâmica calórica, *é* possível, através de um esforço muito grande, conseguir um arranjo, embora seja improvável que você vá atrair discípulos. Contudo, o custo *é* extremamente alto, em termos de coerência com o resto das teorias científicas e com outras observações. Mas por que valeria a pena pagar o preço? Talvez uma pessoa obstinada resolva manter intocada a intuição de que o calor “*é o que ele é e não uma outra coisa.*” Olhando retrospectivamente e sabendo o que sabemos agora, a idéia de que alguém defenderia com unhas e dentes a “intuição do calor” parece uma perda de tempo.

No caso em questão, eu estou predizendo que o poder explicativo, a coerência e a economia favorecerão a hipótese de que a consciência *é* apenas algum padrão de atividade de neurônios. Pode ser que eu esteja

errada. Se eu o estiver, não será porque uma intuição baseada na introspecção é imutável, mas sim porque a ciência nos leva numa direção diferente. Se eu estiver certa, e certos padrões de atividade cerebral *forem* a realidade por trás da experiência, este fato em si não muda minha experiência, de forma que eu (meu cérebro) passe repentinamente a ver meu cérebro como um scanner de ressonância magnética ou um neurocirurgião poderia vê-lo. Eu continuarei a ter experiências da forma habitual, embora meu cérebro precise ter várias experiências e se submeter a muita aprendizagem, de forma que eu possa entender a realidade neuronal delas.

Finalmente, tentando impedir o avanço do dualista, a idéia de que tem que haver uma base de “aparência” subjetiva, sobre a qual as descobertas relativas à realidade/aparência devem repousar é ligeiramente estranha. Parece um pouco como insistir que “em baixo” não pode ser relativo à posição espacial que alguém ocupa; em baixo é em baixo. Ou como insistir que o tempo não pode ser relativo, que ou dois eventos acontecem ao mesmo tempo ou não. Os seres humanos são produtos da evolução; os sistemas nervosos desenvolveram-se dentro do contexto de competição pela sobrevivência – na luta para obter comida, fugir do perigo, lutar e reproduzir. O modelo que o cérebro tem do mundo exterior se aperfeiçoa através da avaliação das várias distinções entre realidade e aparência – em resumo, através da razão crítica comum e da ciência. Na natureza das coisas, é bem provável que o modelo que o cérebro tem de seu mundo interno também permita descobertas relativas à aparência e à realidade. O cérebro não se desenvolveu para conhecer a natureza do sol como um físico a conhece e nem para conhecer a si próprio como conhece o neurofisiologista. Porém, nas devidas circunstâncias, ele poderá de algum modo conhecê-las.<sup>14</sup>

### *O problema está além de nossa débil inteligência*

Inicialmente, essa afirmação parece ser um modesto reconhecimento de nossas limitações. Na verdade, ela é uma poderosa predição baseada não em sólidas evidências, mas numa profunda ignorância (McGuinn 1990). Por tudo o que nós sabemos, a predição pode estar correta, mas pode, da mesma forma, ser muito bem falsa. Quão débil é a nossa inteligência? Qual é o grau de dificuldade do problema? Como seria possível saber que a resolução do problema está além de nosso alcance, independentemente do grau de desenvolvimento da ciência e da tecnologia? Enquanto não se sabe que o cérebro é mais complicado do que inteligente, desistir de tentar descobrir como ele funciona seria desapontador. Por outro lado, enquanto os experimentos continuarem a produzir resultados que contribuam para o nosso entendimento, por que não continuar buscando?<sup>15</sup>

## Rastreamento dos Mecanismos Neurais da Consciência

### Encontrando uma via de acesso

Na neurociência, existem muitos dados de níveis superiores, relevantes para a consciência. Os fenômenos de **blindsight**, **hemineglect**, cérebros comissurotomizados (**split brains**) e anosognosia (inconsciência de déficit) representam grandes restrições para a reflexão teórica. Alguns estudos cuidadosos, utilizando a ressonância magnética nuclear (RMI) e a tomografia por emissão de pósitrons (TEP), têm nos permitido ligar tipos específicos de perdas funcionais a regiões particulares do cérebro.<sup>16</sup> Isso nos ajuda a restringir a série de estruturas a considerar na seleção para uma micro exploração preliminar.

Por exemplo, o hipocampo pode ter parecido um provável candidato a exercer um papel central na consciência porque ele é uma região de tremenda convergência de fibras de diversas áreas do cérebro. Entretanto, nós agora sabemos que a perda bilateral do hipocampo, embora danifique a capacidade de aprender novas coisas, não implica em perda de consciência. No estágio em que estamos, conseguir descartar alguma coisa é em si mesmo um valioso avanço. Nós sabemos também que certas estruturas do tronco cerebral, como o locus cereleus (LC), são indiretamente necessárias, mas não são parte do mecanismo da consciência. O LC desempenha um papel não-específico no surgimento da consciência, mas não um papel específico na consciência de conteúdos particulares, como a consciência, num dado momento, da cor do céu matutino ao invés do som dos regadores da grama. Os dados podem ser fascinantes, mas a questão permanece: como nós podemos passar de um conjunto de dados intrigantes para explicações genuínas do mecanismo básico? Como nós podemos começar?

Ao pensar sobre esse problema, eu tenho recebido uma grande influência de Francis Crick. Sua abordagem básica é simples: se nós vamos resolver o problema, nós devemos tratá-lo como um problema científico, que vai ser enfrentado do mesmo modo como nós enfrentamos outros problemas científicos difíceis. Como qualquer outro mistério científico, o que nós queremos é uma pista experimental relevante. Nós queremos encontrar um fio que, quando puxado, irá revelar muito mais. Para obter isso, nós precisamos formular hipóteses testáveis, que possam conectar macroefeitos com microdinâmica.

O problema que enfrentamos é o seguinte: encontrar fenômenos psicológicos que (a) tenham sido razoavelmente bem estudados pela psicologia experimental; (b) estejam relacionados a dados de lesões de pacientes humanos e a dados de precisas microlesões animais; (c) estejam relacionados a regiões cerebrais bem conhecidas pela neuroanatomia e pela neurofisiologia e (d) cuja conectividade com outras regiões cerebrais esteja

bem estabelecida. A hipótese de trabalho é que se uma pessoa está consciente de um estímulo, seu cérebro será diferente em algum aspecto, quando comparado com a condição em que ela está acordada e atenta, mas inconsciente do estímulo. Uma estratégia promissora é caçar essas diferenças, tendo como guia dados de estudos de lesões, da TEP, de estudos magnetoencefalográficos (MEG), e assim por diante. A descoberta dessas diferenças, em termos de dados neurobiológicos gerais, deve nos ajudar a descobrir uma teoria do mecanismo.

A idéia central é criar uma teoria limitada por dados referentes a vários níveis da organização cerebral – suficientemente limitada, de forma que ela possa ser significativamente testada. Finalmente, uma teoria da consciência terá que abranger uma série de processos envolvidos na consciência, incluindo a atenção e a memória de curto prazo. Inicialmente, entretanto, ela pode visar um subconjunto, como a integração no espaço e no tempo. Se a teoria for falsificada ou se ela sobreviver a duros testes, em ambos os casos nós aprenderemos algo. Ou seja, ou nós teremos descartado possibilidades específicas – o que já é uma boa coisa nesses estágios iniciais de compreensão – ou nós poderemos aprofundar e desenvolver ainda mais a teoria – o que seria ainda melhor. De qualquer modo, o truque é gerar hipóteses substanciais e testáveis ao invés de hipóteses soltas e sem consistência, suscetíveis apenas a experimentos fantasiosos. O truque é fazer algum progresso real.

### Consciência visual

Quais são os candidatos que surgem, dados os limites acima mencionados? As escolhas são bem limitadas. Embora a metacognição, a introspecção e a consciência de emoções, por exemplo, sejam de fato aspectos da consciência, ou nós não temos bons dados relativos a lesões para restringir o espaço da busca de regiões cerebrais relevantes ou a psicofísica é limitada, ou ambos. Consequentemente, é melhor deixar esses processos para estudos posteriores.

A consciência visual, ao contrário, é um candidato mais promissor. No caso da visão, como Crick aponta, há uma vasta literatura sobre psicofísica visual à disposição e há também uma rica literatura de estudos de lesões animais e humanas. Além disso, no que diz respeito ao resto do cérebro, muito já se sabe sobre a neuroanatomia e a neurofisiologia do sistema visual, pelo menos do macaco e do gato. Alguns fenômenos visuais, tais como o “**filling in**”, a competição binocular (**binocular rivalry**), a visão de movimento e a visão de profundidade estereotípica, podem ser ideais para a busca de diferenças neurobiológicas entre o estar e o não estar consciente na vigília. Esse pode ser um bom começo e eu gostaria de enfatizar *começo*.

### A hipótese de Crick

Estando imerso no rico contexto dos detalhes entre os vários níveis, Crick esboçou uma hipótese relativa às estruturas neurais que, segundo ele, tornam as diferenças salientes, dependendo do animal estar ou não visualmente consciente do estímulo.<sup>17</sup> Supõe-se que a integração de representações através de redes neurais espacialmente distribuídas – a unidade na apercepção, por assim dizer – é realizada por uma “ligação” (**binding**) temporal, a saber, uma sincronia nas respostas dos neurônios relevantes. De um modo muito geral, a sugestão de Crick é a seguinte: (1) os córtices primitivos (áreas visuais V1, V2; áreas somatosensoriais S1, S2, etc.) são centrais para a consciência sensorial, como, por exemplo, a consciência visual. Isso se encaixa com os dados relativos a lesões, com dados recentes da TEP (Kosslyn et al. 1993) e também com dados de células simples (**single cell**) (Logothetis e Schall 1989); (2) dentro das primeiras áreas corticais sensoriais, as células piramidais da camada 5 e possivelmente da camada 6 desempenham o papel central.

Até que ponto essa é uma boa idéia? Parte de sua atração deve-se ao fato dela apoiar-se em estruturas básicas. Na biologia, a solução para problemas difíceis sobre o mecanismo podem ser enormemente facilitadas pela identificação de estruturas críticas. Grosseiramente falando, se você sabe “o que”, ajuda muito a imaginar “como”. A hipótese de Crick pode ser apenas uma pequena peça do quebra-cabeça. Se nós tivermos sorte, entretanto, ela ou algo semelhante a ela pode ser uma peça *central* do quebra-cabeça. Mas aqui não é lugar para uma discussão mais detalhada da hipótese de Crick. É suficiente dizer que, verdadeira ou falsa, ela nos fornece uma ousada ilustração de como abordar um problema tão complicado que é frequentemente considerado intratável.

### A hipótese de Llinás

Uma outra alternativa promissora é sugerida pelas diferenças – fenomenológicas e neurobiológicas – entre estados de sono/sonho/vigília (SSV).<sup>18</sup> Essa via de acesso é atrativa, em primeiro lugar, porque existe a familiar e dramática perda de consciência no sono profundo, que é recuperada quando despertamos, e está presente provavelmente também durante o sonho. O fenômeno é visível em muitos sujeitos diferentes e entre várias espécies. Em segundo lugar, as técnicas magnetoencefalográficas e eletroencefalográficas revelam características cerebrais globais, específicas para diferentes estados. Os dados oriundos de lesões em humanos e animais são importantes, especialmente no que concerne a déficits de consciência durante a vigília. Novamente eu percebo a importância das pesquisas sobre **blindsight**, **hemineglect** (tendência a estar inconsciente de estímulos em várias modalidades no lado esquerdo do corpo), agnosia simultânea (incapacidade de ver várias coisas simultaneamente) e anosognosia

(inconsciência de déficits como a paralisia, a cegueira, o discurso atrapalhado, etc.).

Em terceiro lugar, nós temos aprendido muita coisa com as anormalidades e a manipulação do ciclo SSV, juntamente com sua ligação a propriedades cerebrais específicas. Em quarto lugar, algumas das mudanças globais de estado no ciclo SSV, vistas por macrotécnicas, têm sido relacionadas, através de microtécnicas, a interações entre circuitos específicos do córtex e circuitos subcorticais, especialmente circuitos em várias estruturas centrais do tálamo. Em quinto lugar, e mais especificamente, alguns dados magnetoencefalográficos revelam uma robusta forma de onda de 40 Hz durante a vigília e o sono.<sup>19</sup> A definição e a amplitude é muito atenuada durante o sono e a amplitude é modulada durante a vigília e o sono. A análise magnetoencefalográfica da forma da onda nos revela que trata-se de uma onda que se move na direção ântero-posterior do cérebro, cobrindo a distância num tempo de aproximadamente 12 a 13 *milisegundos*. Alguns dados celulares sugerem que essas propriedades dinâmicas emergem de circuitos neurais particulares e de suas propriedades dinâmicas.

Como tudo isso se relaciona? Com base nesses dados e cientes dos vários dados de níveis superiores, Rodolfo Llinás e colaboradores (1991, 1993) levantaram a hipótese de que a organização fundamental subjacente à consciência e as mudanças observadas no padrão SSV são pares de osciladores conectados, que se conectam ambos ao tálamo e ao córtex. No entanto, cada um deles se conecta a populações celulares distintas, de acordo com seu próprio estilo de conectividade particular (figura 6.2). Uma “família” de osciladores conecta os neurônios numa estrutura talâmica conhecida como núcleo intralaminar – uma estrutura com forma de **BAGEL**, cujos neurônios chegam até as camadas superiores do córtex para fornecer uma espécie de ventilação altamente regular para toda a superfície cortical. A outra “família” de osciladores conecta os neurônios aos núcleos talâmicos de informação específica de modalidade (núcleos MS), originando na retina ou na cóclea, por exemplo, áreas corticais especializadas em modalidade (p. ex. V2, S2). Durante o sono profundo, os neurônios intralaminares, que se projetam para o córtex, cessam seu comportamento de 40 Hz.

Durante o sono profundo e o sono, os sinais externos ao córtex são canalizados pelo núcleo reticular do tálamo.

Em termos gerais, a idéia é que a segunda “família” de osciladores fornece o conteúdo (visual, somatosensorial, etc.), enquanto que a primeira fornece o contexto de integração. No sono profundo, os osciladores encontram-se desconectados; no sono eles estão conectados, mas o circuito de oscilação MS não responde a sinais externos da periferia; na

vigília, os osciladores estão conectados e o circuito MS responde aos sinais externos.

Quais são os efeitos de lesões na estrutura talâmica intralaminar? O principal perfil de pequenas lesões unilaterais é a negligência (inconsciência) de todos os estímulos que se originam no lado oposto do corpo. Lesões bilaterais resultam em “inanição”, ou seja, o paciente não inicia nenhum comportamento e quase não responde a estímulos sensoriais ou a questões. Alguns estudos animais mostram basicamente o mesmo perfil.

Por outro lado, lesões nas regiões específicas de modalidade do tálamo ocasionam perdas específicas de modalidade na consciência – a consciência visual, por exemplo, será perdida, mas a consciência de sons, toques, etc. pode permanecer normal. Curiosamente, os magnetoencefalogramas (MEGs) de pacientes com a doença de Alzheimer, que chegam a um estado de degeneração próximo à inanição, apresentam uma forma de onda de 40 Hz, quando ela existe. Esses dados obviamente não são decisivos, mas ao menos são consistentes com a hipótese.

As hipóteses de Llinás e de Crick podem ser combinadas? No mínimo, elas são consistentes. Além disso, elas se apoiam mutuamente nos níveis dos neurônios e das redes. Há ainda um ponto promissor: as duas famílias de osciladores (MS e intralaminar) interconectam-se principalmente na *camada cortical 5* (ver figura 6.2). Do que nós podemos dizer atualmente, essas conexões parecem ser o principal meio pelo qual os osciladores estão conectados. A possibilidade levantada aqui é que a sincronia temporal dos neurônios que carregam sinais sobre estímulos externos, hipotetizada por Crick, pode ser comandada pelo circuito cortical intralaminar. As conexões entre as estruturas do tronco cerebral e o núcleo intralaminar poderiam desempenhar um papel na modulação da excitação e do estado de alerta.

Muitas questões aparecem agora. Por exemplo, como as estruturas centrais da consciência se relacionam com o comportamento? (Ou, como Dennett perguntaria, “o que acontece em seguida?”)<sup>20</sup> Mais especificamente, quais são as conexões entre o núcleo intralaminar e as estruturas motores, e entre a camada 5 dos córtices sensoriais e as estruturas motores? As projeções do núcleo intralaminar ao córtex cingulado desempenham um papel na atenção? Essas são questões motivadas por dados independentes. A convergência de hipóteses serve, é claro, como estímulo. No entanto, é bom lembrar que ela pode também nos remeter aos conselhos proverbiais. A sabedoria nos aconselha a manter um otimismo moderado.

### Observações Conclusivas

Quando se olha as questões pelo lado misterioso dos fenômenos, as soluções podem parecer impossíveis e talvez até mesmo indesejáveis. Do lado da compreensão, entretanto, as soluções parecem quase óbvias e difíceis de serem perdidas de vista. Por que, poder-se-ia imaginar, demorou tanto tempo para se conceber quais são os elementos? Como alguém tão brilhante como Aristóteles pôde deixar de perceber a plausibilidade da idéia de Aristarco de que a Terra era uma esfera que se move em torno do sol? As verdades mais profundas são muito fáceis de serem despercebidas, é claro, assim como é muito fácil para nós despercebermos o que quer que explique por que os animais dormem e sonham, e o que é o autismo. Os problemas são difíceis para a neurociência e para a psicologia experimental, mas à medida que vamos caminhando e que novas técnicas aumentam o acesso não-invasivo aos processos cerebrais globais de seres humanos, as intuições mudam. O que hoje nos parece óbvio foi uma grande novidade para a geração anterior; o que parece confuso para a nossa imaginação é tranqüilamente absorvido pela nova geração de estudantes de pós-graduação. Quem pode nos dizer com certeza se todas as nossas questões sobre a consciência podem ou não ser eventualmente respondidas? Enquanto isso, é recompensador ver algum progresso – ver algumas questões mudarem seu status de Mistérios Que Só Podemos Contemplar com Espanto para Problemas Delicados Que Estamos Começando a Decifrar.

### Referências Bibliográficas

- BICKLE, J. (1992). “Revisionary physicalism”. *Biology and Philosophy*, 7: 411-430
- CHURCHLAND, P. M. (1988). *Matter and consciousness*, 2<sup>nd</sup> Edition. Cambridge, Mass.: MIT Press
- \_\_\_\_\_. (1993a). “Betty Crocker’s theory of the mind: a review of *The Rediscovery of the Mind*, by John Searle”. *London Review of Books* (1995)
- \_\_\_\_\_. (1993b). “Evaluating our self-conception”. *Mind and Language*. (In press).
- CHURCHLAND, P. M. and P. S. CHURCHLAND (1990). “Intertheoretic reduction: a neuroscientist’s field guide”. *Seminars in the Neurosciences*, 4: 249-256
- CHURCHLAND, P. S. (1986). *Neurophilosophy*. Cambridge, Mass.: MIT Press
- \_\_\_\_\_. (1987). “Replies to comments. Symposium on Patricia Smith Churchland’s *Neurophilosophy*”. *Inquiry*, 29: 241-272

- \_\_\_\_\_. (1988). "Reduction and the neurobiological basis of consciousness." In: *Consciousness in contemporary science*. Ed. A.J. Marcel and E. Bisiach. 273-304
- CHURCHLAND, P. S. and T. J. SEJNOWSKI. (1989). "Brain and cognition". In: *Foundations of cognitive science*. Ed. M. Posner. Cambridge, Mass.: MIT Press, 245-300
- \_\_\_\_\_. (1992). *The computational brain*. Cambridge, Mass.: MIT Press
- CRICK, F. C. (1994). *The astonishing hypothesis*. New York: Scribner and Sons
- CRICK, F. C. and C. KOCH (1990). "Towards a neurobiological theory of consciousness". *Seminars in the Neurosciences*, 4: 263-276
- DAMASIO, A. R. (1994). *Descartes' error*. New York: Simon and Schuster
- DAMASIO, H. (1991). "Neuroanatomy of frontal lobe in vivo: A comment on methodology." In: *Frontal Lobe Function and Dysfunction*. Ed. H. Levin, H. Eisenberg and A. Benton. New York: Oxford University Press, 92-121
- DAMASIO, H. and A. R. DAMASIO (1990). The neural basis of memory, language and behavioral guidance: advances with the lesion method in humans. *Seminars in the Neurosciences*, 4: 277-286
- DENNETT, D. (1991). *Consciousness explained*. Boston: Little, Brown and Co.
- FARAH, M. J. (1993). "Neuropsychological inference with an interactive brain: A critique of the 'locality assumption'." *Behavioral and Brain Sciences*. (In Press)
- FEYERABEND, P. K. (1981). *Philosophical papers*, vols. 1 e 2. Cambridge: Cambridge University Press
- FLANAGAN, O. (1992). *Consciousness reconsidered*. Cambridge, Mass.: MIT Press
- \_\_\_\_\_. (1996). "Prospects for a unified theory of consciousness, or, what dreams are made of." In: *Scientific approaches to the question of consciousness: 25<sup>th</sup> Carnegie Symposium on Cognition*. Ed. J. Cohen and J. Schooler. Hillsdale, N.J.: Erlbaum.
- KOSSLYN, S. M., N. M. ALPERT, W. L. THOMPSON, V. MALJKOVIC, S. B. WEISE, C. F. CHABRIS, , S. E. HAMILTON, S. L. RAUCH & S. F. BUONANO (1993). "Visual mental imagery activated topographically organized visual cortex: PET investigations." *Journal of Cognitive Neuroscience*, 5: 263-287
- LLINÁS, R. & D. PARÉ (1991). "Of dreaming and wakefulness." *Neuroscience*, 44: 521-535

- LLINÁS, R. & U. RIBARY (1993). “Coherent 40-Hz oscillation characterizes dream state in humans.” *Proceedings of the National Academy of Sciences*, 90: 2078-2081
- LOGOTHETIS, N & J. D. SCHALL (1989). “Neural correlates of subjective visual perception.” *Science*, 245: 753-761
- LYCAN, W. (1987). *Consciousness*. Cambridge, Mass.: MIT Press
- McGINN, C. (1990). *The problem of consciousness*. Oxford: Blackwells
- SCHAFFNER, K. F. (1993). “Theory structure, reduction, and disciplinary integration in biology.” *Biology and Philosophy*, 8: 319-348
- SEARLE, J. (1992). *The rediscovery of the mind*. Cambridge, Mass.: MIT Press

---

## Notas

Conferência presidencial proferida na ocasião da 67ª Reunião Anual da American Philosophical Association, divisão do Pacífico, em São Francisco, Califórnia, no dia 26 de Março de 1993

<sup>1</sup> Ver a nossa discussão em *The Computational Brain*, Churchland and Sejnowski (1992).

<sup>2</sup> Em relação a opiniões concordantes, ver também Francis Crick (1994); Paul Churchland (1989); Daniel Dennett (1991); Owen Flanagan (1992); William Lycan (1987); John Searle (1992).

<sup>3</sup> Como P. S. Churchland e T. J. Sejnowski argumentaram em (1989).

<sup>4</sup> Para uma excelente discussão do reducionismo, que inclui muitas das complexidades com as quais eu não estou me preocupando muito neste momento, ver Schaffner (1993).

<sup>5</sup> P. S. Churchland (1986), *Neurophilosophy*.

<sup>6</sup> Ver Churchland e Sejnowski (1992); Paul M. Churchland (1993 a).

<sup>7</sup> Ou, o que nós temos preferido mas decidimos não dizer, “o que torna o materialismo revisionista *revisionista*” (P. S. Churchland 1987). Ver também P. M. Churchland (1993). Para uma caracterização parecida, mas de certa forma diferente, ver Bickle (1992).

<sup>8</sup> Ibid. Ver também P. M. Churchland e P. S. Churchland (1990).

<sup>9</sup> Por exemplo, Colin McGinn (1990).

<sup>10</sup> Ver, por exemplo, Feyerabend (1981).

<sup>11</sup> Ver também Owen Flanagan (1992, 1995).

<sup>12</sup> Na seguinte discussão, as idéias são, em sua maior parte, de Paul Churchland (1993a). Para essa discussão, ver “Betty Crocker’s Theory of Mind: A Review of John Searle’s *The Rediscovery of the Mind*”, *London Review of Books* (1995).

<sup>13</sup> Paul Churchland fez essa descoberta em nossa cozinha, há cerca de oito anos atrás. Pareceu-nos um caso típico de alguém que não entende realmente a explicação científica. Ao invés de utilizar adequadamente a teoria da termodinâmica, Betty Crocker mistura-a desajeitadamente com uma velha concepção, como se a velha concepção não precisasse ser modificada. Alguém que pensasse que a eletricidade fosse *causada* por elétrons em movimento, contaria uma história comparável à de Betty Crocker: “a voltagem força os elétrons a se moverem através do fio e, assim fazendo, eles produzem um aumento na eletricidade estática, de modo que as faíscas passam a pular de elétron a elétron, por todo o fio.” Quando eu entretenho audiências de cientistas com a explicação ‘microondas’ de Betty, os risos tornam-se audíveis.

<sup>14</sup> Ver Paul Churchland (1993b).

<sup>15</sup> Ver a discussão convincente e mais detalhada de Daniel Dennett sobre a recusa de McGuinn (Dennett 1991)

<sup>16</sup> Ver especialmente H. Damasio e A. R. Damasio (1990); H. Damasio (1991); A. R. Damasio (1994); Farah (1993).

<sup>17</sup> Essa hipótese é levantada em Crick e Koch (1990) e em Crick (1994).

<sup>18</sup> Ver também minha discussão em P. S. Churchland (1987).

---

<sup>19</sup> Ver Llinás e Paré (1991).

<sup>20</sup> Dennett (1991)