

Consciência*

Autor: Robert Van Gulick

Tradutores: Marco Aurélio & Marcelo Fischborn

Revisor: André Abath

Talvez nenhum aspecto da mente seja mais familiar ou mais enigmático que a consciência e a experiência consciente que temos de nós mesmos e do mundo. O problema da consciência é provavelmente a questão central na teorização atual sobre a mente. Apesar de não haver qualquer acordo nas teorias da consciência, há um consenso difundido, se não universal, de que uma abordagem adequada da mente exige uma compreensão clara da consciência e do seu lugar na natureza. Precisamos entender não apenas o que a consciência é, mas também como se relaciona com outros aspectos, não conscientes, da realidade.

1. História da questão

É provável que questões sobre a natureza da consciência [*conscious awareness*] sejam levantadas desde tanto tempo quanto têm existido os seres humanos. As práticas de sepultamento neolíticas parecem expressar crenças espirituais, e dão indícios primitivos de pensamento ao menos minimamente reflexivo sobre a natureza da consciência humana (Pearson 1999, Clark and Riel-Salvatore 2001). De modo similar, tem-se descoberto que culturas não-letradas adotam invariavelmente alguma forma de concepção espiritual, ou ao menos animista, que indica certo grau de reflexão sobre a natureza da consciência.

Não obstante, alguns defenderam que a consciência, tal como a conhecemos hoje, é um desenvolvimento histórico relativamente recente, que surgiu em algum momento posterior ao período homérico (Jeynes 1974). De acordo com essa concepção, os seres humanos

* Tradução do verbete "Consciousness" de Robert Van Gulick (2004). In: *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.) URL = <<http://plato.stanford.edu/archives/sum2011/entries/consciousness/>>. Publicada com a autorização do editor.

anteriores, incluindo aqueles que lutaram na Guerra de Tróia, não experienciaram a si mesmos como sujeitos internos unificados de seus próprios pensamentos e ações, ao menos não do modo como o fazemos hoje. Outros alegaram que, mesmo durante o período clássico, não havia nenhuma palavra do grego antigo que correspondesse a "consciência" (Wilkes 1984, 1988, 1995). Embora os antigos tivessem muito a dizer sobre questões da mente, é menos claro que tivessem quaisquer conceitos ou preocupações sobre o que entendemos hoje por consciência.

Embora as palavras "consciente" [*conscious*] e "consciência" [*conscience*] sejam usadas de modo bem diferente hoje em dia, é provável que a ênfase da Reforma na última, como uma fonte interna de verdade, tenha desempenhado algum papel na interiorização tão característica da concepção moderna do eu. O Hamlet que andava pelos palcos em 1600 já via o seu mundo e a si mesmo com olhos profundamente modernos.

Aproximadamente no início da era moderna, no século XVII, a consciência alcançou plenamente o centro do pensamento sobre a mente. De fato, da metade do século XVII ao final do século XIX, a consciência foi amplamente tratada como essencial ou definidora do mental. René Descartes definiu a própria noção de pensamento (*pensée*) em termos de consciência reflexiva ou autoconsciência. Nos *Princípios da Filosofia* (1640), ele escreveu:

Pela palavra 'pensamento' (*pensée*), entendo tudo aquilo de que estamos conscientes como operando em nós.

Mais tarde, perto do final do século XVII, John Locke fazia uma afirmação similar, embora um pouco mais qualificada, no *Ensaio sobre o Entendimento Humano* (1688):

Não digo que não há alma em um homem porque ele não está ciente [*sensible*] dela durante o sono. Mas digo que em nenhum momento, acordado ou dormindo, ele pode pensar sem estar ciente disso. Estar ciente não é necessário para nada, exceto nossos pensamentos, e para eles isso é, e sempre será, necessário.

Locke recusava-se explicitamente a fazer qualquer hipótese sobre a base substancial da consciência e sua relação com a matéria, mas tratou-a claramente como essencial ao pensamento, bem como à identidade pessoal.

G. W. Leibniz, contemporâneo de Locke, possivelmente inspirado por seu trabalho matemático sobre diferenciação e integração, ofereceu uma teoria da mente no *Discurso de Metafísica* (1686), que permitia infinitos graus de consciência, e talvez mesmo de pensamentos inconscientes, as chamadas “pequenas percepções” [*petite perceptions*]. Leibniz foi o primeiro a distinguir explicitamente entre percepção e apercepção, isto é, grosseiramente, entre consciência e autoconsciência. Na *Monadologia* (1720), ele também ofereceu sua famosa analogia do moinho, para expressar sua crença de que a consciência não poderia surgir da mera matéria. Ele convidou seu leitor a imaginar alguém caminhando pelo interior de um cérebro expandido, tal como caminharia no interior de um moinho, observando todas as suas operações mecânicas, o que para Leibniz esgotava sua natureza física. Em nenhum lugar, afirma ele, esse observador iria ver quaisquer pensamentos conscientes.

Apesar do reconhecimento de Leibniz da possibilidade de pensamento inconsciente, na maior parte dos dois séculos seguintes os domínios do pensamento e da consciência foram tratados como mais ou menos o mesmo. A psicologia associacionista, seja a buscada por Locke, ou depois no século XVIII por David Hume (1739), ou no século XIX por James Mill (1829), visava descobrir os princípios pelos quais os pensamentos ou ideias conscientes interagiam ou afetavam um ao outro. John Stuart Mill, filho de James Mill, continuou o trabalho de seu pai na psicologia associacionista, mas permitiu que combinações de ideias pudessem produzir resultantes que estivessem para além de suas partes mentais constituintes, formulando assim um modelo primitivo da emergência do mental (1865).

A abordagem puramente associacionista foi criticada no fim do século XVIII por Immanuel Kant (1787), que argumentou que uma teoria adequada da experiência e da consciência fenomenal exigia uma estrutura muito mais rica da organização mental e

intencional. A consciência fenomenal, segundo Kant, não poderia ser uma mera sucessão de ideias associadas, mas teria de ser no mínimo a experiência de um eu consciente situado em um mundo objetivo estruturado com respeito a espaço, tempo e causalidade.

No mundo anglo-americano as abordagens associacionistas continuaram a ser bastante influentes no século XX, tanto na filosofia como na psicologia, enquanto que na esfera germânica e europeia houve um maior interesse na estrutura mais ampla da experiência, que em parte levou ao estudo da fenomenologia, através do trabalho de Edmund Husserl (1913, 1929), Martin Heidegger (1927), Maurice Merleau-Ponty (1945) e outros que expandiram o estudo da consciência aos reinos social, corporal e interpessoal.

No começo da psicologia científica moderna, na metade do século XIX, a mente ainda era amplamente igualada à consciência e os métodos introspectivos dominavam o campo, como no trabalho de Wilhelm Wundt (1897), Hermann von Helmholtz (1897), William James (1890) e Alfred Titchener (1901). Entretanto, a relação entre consciência e cérebro permanecia um grande mistério, como expresso na observação famosa de T. H. Huxley:

Como algo tão notável quanto um estado de consciência resulta da irritação de tecidos nervosos é simplesmente tão inexplicável quanto o surgimento do gênio, quando Aladim esfregou sua lâmpada (1866).

O início do século vinte viu o eclipse da consciência pela psicologia científica, especialmente nos Estados Unidos, com o surgimento do behaviorismo (Watson 1924, Skinner 1953), embora movimentos como a psicologia Gestalt a tenham mantido em curso como uma questão de interesse científico na Europa (Köhler 1929, Köffka 1935). Nos anos 1960, o domínio do behaviorismo enfraquecia com o surgimento da psicologia cognitiva e sua ênfase no processamento de informação e na criação de modelos de processos mentais internos (Neisser 1965, Gardiner 1985). Entretanto, apesar da ênfase renovada na explicação de capacidades cognitivas como memória, percepção e compreensão da linguagem, a

consciência permaneceu como um tópico enormemente negligenciado por mais algumas décadas.

Nos anos 1980 e 90 houve um grande ressurgimento da pesquisa científica e filosófica sobre a natureza e a base da consciência (Baars 1988, Dennett 1991, Penrose 1989, 1994, Crick 1994, Lycan 1987, 1996, Chalmers 1996). Uma vez que a consciência estava de novo em discussão, houve uma rápida proliferação de pesquisas, com uma enxurrada de livros e artigos, e também o início de revistas especializadas (*The Journal of Consciousness Studies*, *Consciousness and Cognition*, *Psyche*), sociedades profissionais (Association for the Scientific Study of Consciousness - ASSC) e conferências anuais dedicadas exclusivamente à sua investigação (Toward a Science of Consciousness, ASSC).

2. Conceitos de consciência

As palavras "consciente" e "consciência" são termos genéricos que abrangem uma ampla variedade de fenômenos mentais. Ambas são usadas com uma diversidade de significados, e o adjetivo "consciente" é heterogêneo em sua extensão, sendo aplicado tanto a organismos inteiros – consciência da criatura – como a estados e processos mentais particulares – consciência do estado (Rosenthal 1986, Gennaro 1995, Carruthers 2000).

2.1 A consciência da criatura

Um animal, uma pessoa ou outro sistema cognitivo pode ser tratado como consciente em vários sentidos diferentes.

Senciência [*sentience*]. Pode ser consciente no sentido genérico de ser simplesmente uma criatura *senciente*, uma criatura capaz de sentir e responder a seu mundo (Armstrong 1981). Ser consciente nesse sentido pode admitir graus, e exatamente qual tipo de capacidades sensoriais é suficiente pode não estar nitidamente definido. Peixes são conscientes no aspecto relevante? E quanto a camarões e abelhas?

Vigília. Pode-se, ademais, exigir que o organismo esteja efetivamente exercendo essas capacidades, e não meramente que tenha a habilidade ou disposição de exercê-las. Assim, poderíamos considerá-lo consciente apenas se estivesse *acordado e em estado normal de alerta*. Nesse sentido, organismos não seriam considerados conscientes quando adormecidos ou em quaisquer dos níveis mais profundos de coma. Novamente, os limites podem ser nebulosos, e casos intermediários podem estar envolvidos. Por exemplo, está-se consciente no sentido relevante quando se está sonhando, hipnotizado ou em estado de fuga [*fugue state*]?

Autoconsciência. Um terceiro e ainda mais exigente sentido poderia definir as criaturas conscientes como aquelas que estão não apenas conscientes, mas também conscientes de que estão conscientes, tratando assim a consciência como uma forma de *autoconsciência* (Carruthers 2000). A exigência de autoconsciência poderia ser interpretada de maneiras variadas, e de acordo com isso variará quais criaturas se qualificariam como conscientes no sentido relevante. Se tomamo-la como envolvendo autoconsciência conceitual explícita, muitos animais não-humanos e até crianças jovens poderiam não se qualificar como tal, mas se apenas formas implícitas mais rudimentares de autoconsciência são exigidas, então uma ampla variedade de criaturas não linguísticas poderiam ser consideradas autoconscientes.

Como é [what it is like]. O famoso critério “*como é*” de Thomas Nagel (1974) visa capturar uma outra, e talvez mais subjetiva, noção de ser um organismo consciente. De acordo com Nagel, um ser é consciente apenas se há “algo como é” ser essa criatura, isto é, algum modo subjetivo com o qual o mundo pareça ou em que se apresente do ponto de vista mental ou experiencial da criatura. No exemplo de Nagel, morcegos são conscientes porque há algo que é como ser um morcego que experiencia seu mundo através de seus sentidos ecolocalizadores, apesar de nós humanos, do nosso ponto de vista humano, não podermos entender empaticamente como tal modo de consciência é do ponto de vista do próprio morcego.

Sujeito de estados conscientes. Uma quinta alternativa seria definir a noção de um organismo consciente em termos de estados conscientes. Isto é, poderíamos primeiro definir o que torna um estado mental um estado mental consciente, e então definir uma criatura

como sendo consciente em termos da posse desses estados. Nosso conceito de um organismo consciente dependeria, assim, da abordagem específica que damos aos estados conscientes (seção 2.2).

Consciência transitiva. Além de descrever criaturas como conscientes nesses vários sentidos, há também sentidos relacionados em que criaturas são descritas como sendo *conscientes de* várias coisas. Essa distinção é por vezes marcada como uma distinção entre noções *transitivas* e *intransitivas* de consciência, com a primeira envolvendo algum objeto ao qual a consciência é dirigida (Rosenthal 1986).

2.2 A consciência de estados

A noção de um estado mental consciente também tem uma variedade de sentidos distintos, embora talvez inter-relacionados. Há ao menos seis opções principais.

Estados em que se sabe que se está [*states one is aware of*]. Em uma leitura comum, um estado mental consciente é simplesmente um estado mental que se sabe estar nele (Rosenthal 1986, 1996). Estados conscientes, nesse sentido, envolvem uma forma de *meta-mentalidade* ou *meta-intencionalidade*, na medida em que exigem estados mentais que são eles mesmos a respeito de estados mentais. Ter um desejo consciente de uma xícara de café é ter esse desejo e também estar simultaneamente e diretamente consciente de que se tem esse desejo. Pensamentos e desejos inconscientes, nesse sentido, são simplesmente aqueles que temos sem saber [*without being aware*] que temos, quer nossa falta de autoconhecimento resulte de simples falta de atenção ou de causas psicanalíticas mais profundas.

Estados qualitativos. Estados podem também ser tratados como conscientes em um sentido aparentemente muito diferente e mais *qualitativo*. Isto é, poderíamos considerar consciente um estado somente se tem ou envolve propriedades qualitativas ou experienciais do tipo ao qual frequentemente nos referimos como “qualia” ou “qualidades sensoriais primárias” [*raw sensory feels*]. (Ver o verbete por Tye 2009.) A percepção que alguém tem do Merlot que está bebendo ou do tecido que está examinando conta como um estado mental

consciente nesse sentido, porque envolve vários qualia sensoriais, por exemplo, qualia de sabor no caso do vinho e qualia de cores na experiência visual que se tem do pano. Há um considerável desacordo sobre a natureza de tais qualia (Churchland 1985, Shoemaker 1990, Clark 1993, Chalmers 1996) e igualmente sobre sua existência. Tradicionalmente, os qualia foram tratados como características monádicas da experiência, privadas e inefáveis, mas as teorias atuais dos qualia frequentemente rejeitam ao menos algum desses comprometimentos (Dennett 1990).

Estados fenomenais. Por vezes refere-se a esses qualia como propriedades fenomenais e ao tipo associado de consciência como consciência *fenomenal*, mas o último termo talvez seja mais propriamente aplicado à estrutura completa da experiência e envolva muito mais que os qualia sensoriais. A estrutura fenomenal da consciência também abrange muito da organização espacial, temporal e conceitual de nossa experiência do mundo e de nós mesmos como agentes nele (ver seção 4.3). Por isso, é provavelmente melhor, ao menos inicialmente, distinguir o conceito de consciência fenomenal do de consciência qualitativa, embora sem dúvida eles se sobreponham.

Estados como-é [what-it-is-like]. A consciência em ambos esses sentidos liga-se também à noção de Thomas Nagel (1974) de uma criatura consciente, na medida em que se considera um estado mental consciente no sentido de “*como é*” simplesmente se há algo que seja como estar nesse estado. O critério de Nagel poderia ser entendido como buscando fornecer uma concepção em primeira pessoa, ou interna, do que torna um estado um estado fenomenal ou qualitativo.

Consciência de acesso. Os estados poderiam ser conscientes em um sentido de acesso, aparentemente muito diferente, que tem mais a ver com as relações intramentais. A este respeito, um estado ser consciente é uma questão de sua disponibilidade para interagir com outros estados e do acesso que se tem ao seu conteúdo. Nesse sentido mais funcional, que corresponde ao que Ned Block (1995) chama de *consciência de acesso*, um estado visual ser consciente não é tanto uma questão de ter ou não um “*como é*” [*what it's likeness*] qualitativo, mas de ele e a informação visual que carrega estarem ou não geralmente disponíveis para uso

e orientação do organismo. Na medida em que a informação nesse estado está rica e flexivelmente disponível para o organismo que a contém, então ele conta como um estado consciente no aspecto relevante, quer ou não tenha qualquer percepção [*feel*] qualitativa ou fenomenal no sentido de Nagel.

Consciência narrativa. Os estados também poderiam ser considerados conscientes em um sentido *narrativo* que apela à noção de um “fluxo de consciência”, considerado como uma narração de episódios progressiva, mais ou menos serial, da perspectiva de um eu efetivo ou meramente virtual. A ideia seria igualar os estados mentais conscientes da pessoa com aqueles que aparecem no fluxo (Dennett 1991, 1992).

Embora essas seis noções sobre o que torna consciente um estado possam ser especificadas independentemente, elas obviamente não o podem sem conexões potenciais, e nem esgotam o reino das opções possíveis. Extraíndo conexões, poderíamos defender que estados aparecem no fluxo de consciência apenas na medida em que estamos conscientes [*aware*] deles, e assim forjar um vínculo entre a primeira noção meta-mental de um estado consciente e o conceito de um fluxo ou narrativa. Ou poderíamos conectar a noção de acesso com a noção fenomenal ou qualitativa de um estado consciente, tentando mostrar que estados que representam daquele modo tornam seus conteúdos amplamente disponíveis no aspecto exigido pela noção do acesso.

Buscando ir além das seis opções, poderíamos distinguir estados conscientes de não-conscientes apelando a aspectos de sua dinâmica intramental e de suas interações outras que não as meras relações de acesso; por exemplo, estados conscientes poderiam manifestar um estoque mais rico de interações sensíveis ao conteúdo, ou um grau maior de orientação intencional flexível, do tipo associado com o controle autoconsciente do pensamento. De modo alternativo, poderíamos tentar definir estados conscientes em termos de criaturas conscientes. Isto é, poderíamos dar uma explicação [*account*] do que é ser uma criatura consciente ou talvez até mesmo de um eu consciente, e então definir a noção de um estado consciente em termos de ser um estado dessa criatura ou sistema, o que seria o inverso da

última opção considerada acima para definir criaturas conscientes em termos de estados mentais conscientes.

2.3 Consciência como uma entidade

O substantivo “consciência” tem uma variedade igualmente diversificada de significados, que em grande medida são paralelos àqueles do adjetivo “consciente”. Podem ser feitas distinções entre a consciência de criaturas e a de estados, bem como entre as variedades de cada uma. Pode-se referir especificamente à consciência fenomenal, consciência de acesso, consciência reflexiva ou meta-mental e à consciência narrativa, entre outras variedades.

Aqui a consciência em si não é tratada tipicamente como uma entidade substancial, mas meramente como a reificação abstrata de qualquer propriedade ou aspecto que é atribuído pelo uso relevante do adjetivo “consciente”. Consciência de acesso é simplesmente a propriedade de ter o tipo exigido de relações internas de acesso, e a consciência qualitativa é simplesmente a propriedade que é atribuída quando “consciente” é aplicado, no sentido qualitativo, a estados mentais. O quanto isso por si só nos compromete com o estatuto ontológico da consciência dependerá de quão platônico se é a respeito de universais em geral (ver o verbete por Klima 2008). Isso não precisa nos comprometer com a consciência como uma entidade distinta mais do que o uso de “quadrado”, “vermelho” ou “suave” nos compromete com a existência da quadradeza [*squareness*], vermelhidão ou suavidade como entidades distintas.

Embora não seja uma norma, poderíamos ainda assim adotar uma concepção realista mais robusta da consciência como um componente da realidade. Isto é, poderíamos pensar a consciência como algo mais parecido com campos eletromagnéticos do que com a vida.

Desde a morte do vitalismo, não pensamos a vida *por si só* como algo distinto de coisas vivas. Há coisas vivas, incluindo organismos, estados, propriedades e partes de organismos, comunidades e linhagens evolutivas de organismos, mas a vida não é ela mesma mais uma coisa, um componente adicional da realidade, alguma força vital que se acrescenta às coisas

vivas. Aplicamos corretamente os adjetivos “vivo” e “animado” [*alive*] a muitas coisas, e ao fazer isso poderíamos dizer que lhes estamos atribuindo vida, mas sem nenhum significado ou realidade além daquele envolvido em serem coisas vivas.

Campos eletromagnéticos, por contraste, são tratados como partes reais e independentes do nosso mundo físico. Mesmo que por vezes se possa especificar os valores de um certo campo apelando ao comportamento das partículas em seu interior, os campos são tratados eles mesmos como constituintes concretos da realidade, e não meramente como abstrações ou conjuntos de relações entre partículas.

De modo similar, poderíamos tratar “consciência” como se referindo a um componente ou aspecto da realidade que manifesta a si mesmo em estados e criaturas conscientes, mas que é mais do que meramente a nominalização abstrata do adjetivo “consciente” que lhes aplicamos. Embora essas concepções fortemente realistas não sejam muito comuns atualmente, elas devem ser incluídas no espaço lógico de opções.

Há assim muitos conceitos de consciência, e tanto “consciente” como “consciência” são usados de modos muito variados, sem um significado privilegiado ou canônico. No entanto, isso pode ser menos um problema do que uma excessiva fartura. A consciência é uma característica complexa do mundo, e entendê-la exigirá uma diversidade de ferramentas conceituais para lidar com seus vários aspectos diferentes. A pluralidade conceitual é, assim, justamente o que se esperaria. Na medida em que se evita a confusão, sendo claro quanto a seus significados, há um grande valor em se ter uma variedade de conceitos pelos quais podemos acessar e apreender a consciência em toda a sua rica complexidade. No entanto, não se deve supor que a pluralidade conceitual acarreta divergência de referência. Nossos múltiplos conceitos de consciência podem, de fato, apanhar aspectos variados de um mesmo e unificado fenômeno mental subjacente. Se assim o é, e em que medida, permanece uma questão em aberto.

3. Problemas da consciência

A tarefa de entender a consciência é um projeto igualmente diversificado. Não só muitos aspectos diferentes da mente contam como conscientes em algum sentido, mas cada um é também aberto a vários aspectos em que poderia ser explicado ou modelado. Entender a consciência envolve não apenas uma multiplicidade de coisas a serem explicadas [*explananda*], mas também de questões que elas colocam e tipos de respostas que exigem. Sob risco simplificar excessivamente, as questões relevantes podem ser reunidas sob três rubricas cruas como as questões de o que é, como é e de por que existe:

A questão descritiva: *O que é a consciência?* Quais são suas características principais? E por que meios elas podem ser mais bem descobertas, descritas e modeladas?

A questão explicativa: *Como a consciência do tipo relevante passa a existir?* Ela é um aspecto primitivo da realidade e, se não, como a consciência no aspecto relevante surge (ou poderia surgir), ou é causada, por entidades ou processos não-conscientes?

A questão funcional: *Por que a consciência do tipo relevante existe?* Ela tem uma função e, se sim, qual? Ela age causalmente e, se sim, quais seus tipos de efeitos? Ela faz diferença para a operação dos sistemas em que está presente e, se sim, por que e como?

As três questões focam respectivamente em descrever as características da consciência, explicar sua base ou causa subjacente e explicar sua função ou valor. A divisão em três é certamente um tanto artificial e, na prática, as respostas que se dá a cada uma dependerão em parte do que se diz sobre as outras. Não se pode, por exemplo, responder adequadamente à questão do que é a consciência e descrever suas características principais sem abordar a questão do porquê, de seu papel funcional nos sistemas cujas operações afeta. Nem se poderia explicar como o tipo relevante de consciência poderia surgir de processos não-conscientes a menos que se tivesse uma abordagem clara de precisamente quais características teriam de ser causadas ou realizadas para contar como a produzindo. Apesar dessas ressalvas, a divisão

em três vias de questões fornece uma estrutura útil para articular o projeto explicativo global, e para avaliar a adequação de teorias ou modelos particulares da consciência.

4. A questão descritiva: *Quais* são as características da consciência?

A questão do *que* é a consciência nos pede para descrever e modelar as características principais da consciência, mas exatamente quais características são relevantes variará de acordo com o tipo de consciência que pretendermos capturar. As propriedades principais da consciência de acesso podem ser completamente diferentes daquelas da consciência qualitativa ou fenomenal, e as da consciência reflexiva ou narrativa podem diferir de ambas. Entretanto, acumulando teorias detalhadas de cada tipo, podemos ter a esperança de encontrar ligações importantes entre elas, e talvez até descobrir que coincidem em ao menos alguns aspectos-chave.

4.1 Dados de primeira e de terceira pessoa

O projeto descritivo geral exigirá uma variedade de métodos de investigação (Flanagan 1992). Embora pudéssemos tratar ingenuamente os fatos da consciência como bastante autoevidentes para exigir quaisquer métodos sistemáticos de obtenção de dados, a tarefa epistêmica está realmente longe de ser trivial (Husserl 1913).

O acesso introspectivo de primeira pessoa proporciona uma fonte de conhecimento [*insight*] de nossa vida mental consciente, mas não é por si só suficiente nem mesmo útil, a menos que usado de modo treinado e disciplinado. Obter os indícios necessários da estrutura da experiência exige tanto que nos tornemos observadores fenomenologicamente sofisticados quanto que complementemos nossos resultados introspectivos com vários tipos de dados de terceira pessoa disponíveis para o observador externo (Searle 1992, Varela 1995, Siewert 1998).

Como os fenomenólogos sabem há mais de um século, descobrir a estrutura da experiência consciente demanda um rigoroso olhar introspectivo [*inner directed stance*] que é completamente diferente de nossa forma cotidiana de autoconsciência (Husserl 1929, Merleau-Ponty 1945). A observação habilidosa do tipo necessário exige treino, esforço e a capacidade de adotar perspectivas alternativas sobre a própria experiência.

A necessidade de dados empíricos de terceira pessoa obtidos por observadores externos é talvez mais óbvia com respeito aos tipos claramente mais funcionais de consciência, como a consciência de acesso, mas é exigida mesmo com respeito à consciência fenomenal e qualitativa. Por exemplo, estudos de déficit que correlacionam várias regiões de danos neurais e funcionais com anomalias da experiência consciente podem tornar-nos conscientes de aspectos da estrutura fenomenal que escapam à nossa consciência introspectiva normal. Como esses estudos de caso mostram, coisas que parecem inseparavelmente unidas ou singulares, do nosso ponto de vista normal de primeira pessoa, podem tornar-se separadas na experiência (Sacks 1985, Shallice 1988, Farah 1995).

Ou, para apanhar um outro exemplo, dados de terceira pessoa podem nos tornar conscientes de como nossas experiências de estar agindo e nossas experiências da duração de eventos afetam umas as outras, de um modo que nunca poderíamos discernir através da mera introspecção (Libert 1985, Wegner 2002). E esses fatos obtidos por métodos de terceira pessoa nem são meramente a respeito das causas ou bases da consciência; eles frequentemente dizem respeito à estrutura mesma da própria consciência fenomenal. Métodos interativos de primeira pessoa, terceira pessoa e talvez até de segunda pessoa (Varela 1995) serão todos exigidos para coletar a evidência necessária.

Usando todas essas fontes de dados, estaremos, espera-se, aptos a construir modelos descritivos detalhados dos vários tipos de consciência. Embora características específicas da maior importância possam variar entre os diferentes tipos, nosso projeto descritivo global precisará enfrentar ao menos os sete aspectos gerais da consciência a seguir (seções 4.2-4.8).

4.2 O caráter qualitativo

O *caráter qualitativo* é frequentemente igualado às chamadas “sensações brutas” [*raw feels*], e ilustrado pela vermelhidão que se experiencia quando se olha para tomates maduros ou o sabor doce específico que se encontra ao provar um abacaxi igualmente maduro (Locke 1968). O tipo relevante de caráter qualitativo não está restrito a estados sensoriais, mas é tipicamente tomado como estando presente enquanto um aspecto dos estados experienciais em geral, tais como os pensamentos ou desejos experienciados (Siewert 1998).

Para alguns, a existência dessas sensações parece demarcar o ponto a partir do qual estados ou criaturas são realmente conscientes. Se um organismo sente e responde de maneira adequada a seu mundo, mas carece de tais qualia, então no máximo poderia contar como consciente em um sentido frouxo e não-literal. Ou ao menos assim pareceria àqueles que tomam a consciência qualitativa, no sentido de “como é”, como sendo central filosófica e cientificamente (Nagel 1974, Chalmers 1996).

Problemas de variadas formas sobre os qualia se avultaram no passado recente – Pode haver qualia invertidos? (Block 1980a, 1980b, Shoemaker 1981, 1982) Os qualia são epifenomenais? (Jackson 1982, Chalmers 1996) Como os estados neurais podem originar qualia? (Levine 1983, McGinn 1991) Mas a questão do que é a consciência levanta um problema mais básico sobre os qualia, a saber, o de dar uma descrição clara e articulada de nosso espaço de qualia e do estatuto de seus qualia específicos.

Na ausência de tal modelo, erros factuais ou descritivos são muito prováveis. Por exemplo, afirmações sobre a ininteligibilidade da ligação entre o vermelho experienciado e qualquer substrato neural possível dessa experiência tratam, às vezes, o quale da cor relevante como uma propriedade simples e *sui generis* (Levine 1983), mas a vermelhidão fenomenal existe de fato em um espaço complexo de cores, com múltiplas dimensões sistemáticas e relações de semelhança (Hardin 1992). Entender o quale específico da cor em relação a essa ampla estrutura relacional não nos dá apenas uma melhor apreensão descritiva

da sua natureza qualitativa, mas pode também fornecer-nos alguns “ganchos”, aos quais poderíamos anexar as ligações psicofísicas inteligíveis.

A cor pode ser a exceção quanto a termos um entendimento formal específico e bem desenvolvido do espaço qualitativo relevante, mas provavelmente não é uma exceção com respeito à importância desses espaços para o nosso entendimento das propriedades qualitativas em geral (Clark 1993, P. M. Churchland 1995). (Ver o verbete “qualia”, por Tye 2009)

4.3 A estrutura fenomenal

A *estrutura fenomenal* não deve ser confundida com a estrutura qualitativa, apesar do uso por vezes intercambiável de “qualia” e “propriedades fenomenais” na literatura. A “organização fenomenal” abrange todos os vários tipos de ordem e estrutura encontrados no domínio da experiência, isto é, no domínio do mundo tal como nos *aparece*. Obviamente há ligações importantes entre o fenomenal e o qualitativo. Com efeito, os qualia talvez fossem mais bem entendidos como propriedades dos objetos fenomenais ou experienciados, mas de fato há bem mais para o fenomenal do que sensações brutas [*raw feels*]. Como Kant (1787), Husserl (1913) e gerações de fenomenólogos mostraram, a estrutura fenomenal da experiência é ricamente intencional e envolve não apenas ideias e qualidades sensoriais, mas representações complexas de tempo, espaço, causa, do corpo, de si mesmo, do mundo e da estrutura organizada da realidade vivida, em todas as suas formas conceituais e não-conceituais.

Uma vez que muitos estados não-conscientes também têm aspectos intencionais e representacionais, pode ser melhor considerar a estrutura fenomenal como envolvendo um tipo especial de organização e conteúdo intencional e representacional, o tipo distintivamente associado à consciência (Siewert 1998). (Ver o verbete sobre “teorias representacionais da consciência”, por Lycan 2008).

Responder a questão do que é a consciência exige uma abordagem cuidadosa da estrutura representacional coerente e densamente organizada em que as experiências particulares estão inseridas. Dado que a maior parte dessa estrutura está apenas implícita na organização da experiência, ela não pode simplesmente ser lida por introspecção. Articular a estrutura do domínio fenomenal de modo claro e inteligível é um longo e difícil processo de inferência e construção de modelos (Husserl 1929). A introspecção pode ajudar, mas muita construção teórica e engenhosidade são também necessárias.

4.4 Subjetividade

A *subjetividade* é mais uma noção por vezes igualada na literatura aos aspectos qualitativos ou fenomenais da consciência, mas novamente há boas razões para reconhecê-la, ao menos em algumas de suas formas, como um aspecto distinto da consciência – relacionada ao qualitativo e ao fenomenal, mas diferente de cada um. Em particular, a forma epistêmica da subjetividade diz respeito aos aparentes limites da cognoscibilidade ou mesmo da compreensibilidade de vários fatos sobre a experiência consciente (Nagel 1974, Van Gulick 1985, Lycan 1996).

Na abordagem de Thomas Nagel (1974), os fatos sobre como é ser um morcego são subjetivos no sentido relevante, porque podem ser completamente entendidos apenas a partir do tipo de ponto de vista do morcego. Apenas criaturas capazes de ter ou experienciar de modo similar tais experiências podem entender seu como-é no sentido empático necessário. Fatos sobre a experiência consciente podem ser entendidos no máximo incompletamente de um ponto de vista exterior, de terceira pessoa, tal como os que estão associados com a ciência física objetiva. Uma concepção similar sobre os limites da teoria em terceira pessoa parece estar por traz de afirmações a respeito do que a Maria hipotética de Frank Jackson (1982), a super cientista de cores, não poderia entender sobre experienciar o vermelho, devido à sua própria história empobrecida de experiência visual acromática.

É um debate aberto se os fatos sobre a experiência são realmente limitados epistemicamente dessa maneira (Lycan 1996), mas a alegação de que entender a consciência exige formas especiais de conhecimento e de acesso a partir do ponto de vista interior é intuitivamente plausível e tem uma longa história (Locke 1688). Assim, qualquer resposta adequada à questão do que é a consciência tem de cobrir o estatuto epistêmico da consciência, ambas as nossas habilidades e limites para compreendê-la (Papineau 2002, Chalmers 2003). (Ver o verbete sobre “auto-conhecimento”, por Gertler 2011).

4.5 Organização autoperspectivista

A estrutura perspectivista da consciência é um aspecto de sua organização fenomenal global, mas é importante o suficiente para merecer discussão por si só. Na medida em que a perspectiva chave é aquela do eu consciente, a característica específica poderia ser chamada de *autoperspectividade*. Experiências conscientes não existem como átomos mentais isolados, mas como modos ou estados de um eu ou sujeito consciente (Descartes 1644, Searle 1992, em contraste com Hume 1739). Uma experiência visual de uma esfera azul é sempre uma questão de haver algum eu ou sujeito para o qual pareceu desse modo. Uma dor forte ou aguda é sempre uma dor sentida ou experienciada por algum sujeito consciente. O eu não precisa aparecer como um elemento explícito em nossas experiências, mas, como Kant (1787) notou, o “eu penso” tem de acompanhar ao menos potencialmente cada uma delas.

O eu poderia ser tomado como o ponto de vista perspectivista a partir do qual o mundo de objetos está presente para a experiência (Wittgenstein 1921). Ele fornece não apenas uma perspectiva espacial e temporal para nossa experiência do mundo, mas também de significado e inteligibilidade. A coerência intencional do domínio experiencial repousa sobre a interdependência dual entre o eu e o mundo: o eu como perspectiva a partir da qual os objetos são conhecidos e o mundo como a estrutura integrada de objetos e eventos cujas possibilidades de serem experienciados implicitamente definem a natureza e a localização do eu (Kant 1787, Husserl 1929).

Os organismos conscientes obviamente diferem na medida em que constituem um eu unificado e coerente, e provavelmente diferem quanto ao tipo ou grau de foco perspectivista que incorporam em suas respectivas formas de experiência (Lorenz 1977). A consciência pode não exigir um eu distinto ou substancial do tipo tradicional cartesiano, mas ao menos algum grau de organização perspectivista semelhante a um ego parece essencial para a existência de algo que poderia contar como uma experiência consciente. Parece que as experiências não seriam capazes de existir sem um eu ou sujeito para experienciá-las mais do que poderiam as ondas do oceano sem o mar através do qual se movem. A questão descritiva exige, assim, alguma abordagem do aspecto autoperspectivista da experiência e da organização de tipo egoica [*self-like*] das mentes conscientes da qual depende, mesmo que a abordagem relevante trate o eu de um modo relativamente deflacionário e virtual (Dennett 1991, 1992).

4.6 Unidade

A *unidade* está intimamente ligada à autoperspectiva, mas merece sua própria menção como um aspecto chave da organização da consciência. Sistemas conscientes e estados mentais conscientes envolvem ambos muitas formas diversificadas de unidade. Algumas são unidades causais, associadas à integração da ação e controle em um foco unificado de agência. Outras são formas mais representacionais e intencionais de unidade, envolvendo a integração de diversos itens do conteúdo em ligações de muitas escalas e níveis (Cleeremans 2003).

Algumas dessas integrações são relativamente locais, como quando diversas características detectadas com uma única modalidade sensorial são combinadas na representação de objetos externos portando tais características, por exemplo, quando se tem a experiência visual consciente de uma lata vermelha de sopa passando sobre um guardanapo verde listrado (Triesman and Gelade 1980).

Outras formas de unidade intencional envolvem uma gama muito mais ampla de conteúdos. O conteúdo da experiência atual que alguém tem da sala em que está sentado

depende, em parte, de sua localização em uma estrutura bem maior, associada com a consciência de sua existência como observador temporalmente estendido em um mundo de objetos espacialmente conectados e existentes de modo independente (Kant 1987, Husserl 1913). A experiência individual pode ter o conteúdo que tem apenas porque reside nessa estrutura unificada mais ampla de representação. (Ver o verbete “Unidade da consciência”, por Brook e Raymont 2010.)

4.7 Intencionalidade e transparência

Estados mentais conscientes são tipicamente tratados como tendo um aspecto representacional ou intencional, na medida em que são sobre coisas, referem-se a coisas ou têm condições de satisfação. A experiência visual de alguém *representa corretamente* o mundo se há violetas em um vaso branco sobre a mesa (em oposição a Travis 2004), sua memória consciente é *do* ataque ao World Trade Center e seu desejo consciente é *por* um copo de água gelada. Entretanto, estados não-conscientes também podem exibir intencionalidade desse modo, e é importante entender de que maneiras os aspectos representacionais de estados conscientes se assemelham e diferem dos não-conscientes (Carruthers 2000). Searle (1990) oferece um ponto de vista contrário, de acordo com o qual apenas estados conscientes e disposições para ter estados conscientes podem ser genuinamente intencionais, mas a maioria dos teóricos tratam a intencionalidade como estendendo-se amplamente no interior do domínio inconsciente. (Ver o verbete “Consciência e intencionalidade”, por Siewert 2011).

Uma dimensão potencialmente importante da distinção diz respeito à assim chamada *transparência*, que é uma característica importante da consciência em dois sentidos metafóricos inter-relacionados, cada um dos quais com aspectos intencionais, experienciais e funcionais.

Diz-se frequentemente que a experiência perceptual consciente é transparente ou, na expressão de G. E. Moore (1922), “diáfana”. “Olhamos através” de nossa experiência sensorial transparentemente, na medida em que parecemos diretamente conscientes dos objetos e

eventos externos que presenciamos, em vez de estarmos conscientes de quaisquer propriedades da experiência pelas quais ela nos apresenta ou representa tais objetos. Quando olho o vento a soprar no pasto, é da grama verde ondulante que estou consciente, e não de qualquer propriedade verde da minha experiência visual. (Ver o verbete “Teorias representacionais da consciência”, por Lycan 2008.) O próprio Moore acreditava que poderíamos nos tornar conscientes dessas qualidades posteriores com esforço e redirecionamento da atenção, embora alguns defensores contemporâneos da transparência o neguem (Harman 1990, Tye 1995).

Pensamentos e experiências conscientes também são transparentes em um sentido semântico, em que seus significados nos parecem imediatamente conhecidos no próprio ato de pensá-los (Van Gulick 1992). Nesse sentido, poderíamos dizer que “através deles pensamos diretamente” no que significam ou representam. Nesse sentido semântico, a transparência pode corresponder, ao menos em parte, ao que John Searle chama de “intencionalidade intrínseca” da consciência (Searle 1992).

Nossos estados mentais conscientes parecem ter seus significados intrinsecamente, ou a partir do interior, apenas pelo que são por si mesmos, em contraste com muitas teorias externalistas do conteúdo mental, que baseiam o significado em relações causais, contrafactuais ou informacionais entre seus objetos semânticos ou referenciais.

A concepção de conteúdos conscientes como determinados intrinsecamente e internamente autoevidentes é por vezes apoiada por apelos a intuições sobre o cérebro na cuba, que fazem parecer que os estados mentais conscientes do cérebro encubado manteriam todos os seus conteúdos intencionais normais, apesar da falta de todas as suas ligações causais e informacionais normais com o mundo (Horgan e Tienson 2002). Há permanente controvérsia sobre tais casos e sobre as concepções concorrentes internalistas (Searle 1992) e externalistas (Dretske 1995) da intencionalidade consciente.

Embora a transparência semântica e a intencionalidade intrínseca tenham algumas afinidades, não devem ser simplesmente igualadas, já que é possível acomodar a primeira noção em uma abordagem mais externalista do conteúdo e do significado. Obviamente, tanto a

transparência semântica como a sensorial não apenas dizem respeito aos aspetos representacionais e intencionais da consciência, mas também são aspectos experienciais da nossa vida consciente. Ambas também têm aspectos funcionais, na medida em que as experiências conscientes interagem umas com as outras de maneiras ricamente apropriadas ao conteúdo, que manifestam nosso entendimento transparente de seus conteúdos.

4.8 Fluxo dinâmico

A *dinâmica* da consciência é evidente na ordem coerente de seu processo constante de fluxo e autotransformação, o que William James (1890) chamou de “*fluxo da consciência*”. Algumas sequências temporais da experiência são geradas por fatores puramente internos, como quando se raciocina sobre um quebra-cabeça, e outras dependem em parte de causas externas, como quando se persegue uma bola no ar, mas mesmo as últimas são em grande medida moldadas pelo modo como a consciência transforma a si mesma.

Seja parcialmente em resposta a influências de fora, ou inteiramente do interior, de momento a momento cada sequência da experiência cresce coerentemente para além daquelas que a precederam, constrangida e permitida pela estrutura global de ligações e limites incorporados em sua organização anterior subjacente (Husserl 1913). A esse respeito, a consciência é um sistema autopoietico, isto é, um sistema autocriador e auto-organizador (Varela e Maturana 1980).

Como um agente mental consciente, posso fazer muitas coisas, tais como vasculhar meu quarto, vasculhar uma imagem mental dele, rever na memória os pratos de uma refeição recente em um restaurante, juntamente com seus sabores e aromas, raciocinar passo a passo sobre um problema complexo, ou planejar uma ida às compras e executar o plano quando chegar ao mercado. Todos esses são hábitos e atividades comuns, mas cada um envolve a geração dirigida de experiências, de uma maneira que manifesta uma compreensão prática implícita de suas propriedades intencionais e conteúdos interconectados (Van Gulick 2000).

A consciência é um processo dinâmico, e assim uma resposta descritiva adequada à questão do que ela é tem de lidar com mais do que apenas suas propriedades estáticas ou momentâneas. Em particular, tem de dar alguma explicação da dinâmica temporal da consciência e da maneira pela qual seu fluxo autotransformante reflete tanto sua coerência intencional quanto a autocompreensão semântica incorporada nos controles organizados através dos quais as mentes conscientes refazem continuamente a si mesmas, enquanto sistemas autopoieticos engajados em seus mundos.

Uma abordagem descritiva exaustiva da consciência precisaria lidar com mais do que apenas essas sete características, mas ter uma abordagem clara de cada uma delas seria andar um longo caminho em direção a responder a questão “O que é a consciência?”.

5. A questão explicativa: *Como* a consciência pode existir?

A questão de *como* a consciência pode existir foca na explicação, antes que na descrição. Pede-nos para explicar o estatuto básico da consciência e seu lugar na natureza. Ela é por si só uma característica fundamental da realidade, ou sua existência depende de outros itens, não-conscientes, sejam eles físicos, biológicos, neurais ou computacionais? Se a segunda opção é o caso, podemos explicar ou entender como os itens não-conscientes relevantes puderam causar ou realizar a consciência? Ou, colocado de modo simples, podemos explicar como fazer algo consciente a partir de coisas que não são conscientes?

5.1 A diversidade de projetos explicativos

A questão de como pode existir a consciência não é uma questão única, mas antes uma família geral de questões mais específicas (Van Gulick 1995). Todas dizem respeito à possibilidade de explicar algum gênero ou aspecto da consciência, mas variam no que explicar [*explananda*], na explicação [*explanans*] e nos critérios para uma explicação bem sucedida. Por exemplo, poderíamos perguntar se podemos explicar a consciência de acesso

computacionalmente, simulando as relações de acesso necessárias em um modelo computacional. Ou, em vez disso, poderíamos nos interessar em se as propriedades fenomenais e qualitativas da mente de uma criatura consciente podem ser deduzidas *a priori* de uma descrição das propriedades neurais de seus processos cerebrais. Ambas são versões da questão de como a consciência pode existir, mas perguntam sob perspectivas de projetos explicativos muito diferentes, e assim podem diferir em suas respostas (Lycan 1996). Seria impraticável, se não impossível, catalogar todas as versões possíveis da questão de como a consciência existe, mas algumas das principais opções podem ser listadas.

O que explicar [explananda]. Os possíveis explicáveis incluiriam os vários tipos de consciência de estados e de criaturas distinguidos acima, bem como as sete características da consciência listadas em resposta à questão de o que é a consciência. Esses dois tipos de coisas explicáveis se sobrepõem e interagem. Poderíamos, por exemplo, buscar explicar o aspecto dinâmico tanto da consciência fenomenal como da de acesso. Ou poderíamos tentar explicar a subjetividade tanto da consciência qualitativa como da metamental. Nem toda característica se aplica a cada gênero de consciência, mas todas se aplicam a vários. A maneira de se explicar uma dada característica em relação a um gênero de consciência pode não corresponder ao que é necessário para explicá-la em relação a um outro.

Explicações [explanans]. O espectro de explicações possíveis também é diversificado. Talvez em sua forma mais ampla, a questão de como pode existir a consciência pergunta como a consciência do tipo relevante poderia ser causada ou realizada por itens não-conscientes, mas podemos gerar uma profusão de questões mais específicas, restringindo ainda mais o escopo das explicações relevantes. Poderíamos buscar explicar como uma dada característica da consciência é causada ou realizada por processos *neurais*, estruturas *biológicas*, mecanismos *físicos*, relações *funcionais* ou *teleofuncionais*, por uma organização *computacional* ou mesmo por estados *mentais não-conscientes* subjacentes. As perspectivas para o sucesso explicativo irão variar de acordo. Em geral, quanto mais limitado e elementar o escopo da explicação [explanans], mais difícil é o problema de explicar como isso poderia ser suficiente para produzir a consciência (Van Gulick 1995).

Crítérios de explicação. O terceiro parâmetro chave é como se define o critério para uma explicação bem sucedida. Poderíamos exigir que o que é explicado seja dedutível *a priori* da explicação, embora seja controverso se isso é um critério necessário ou mesmo suficiente para explicar a consciência (Jackson 1993). Sua suficiência dependerá em parte da natureza das premissas a partir das quais a dedução será feita. Como questão de lógica, serão necessários alguns princípios de ponte para conectar proposições ou frases sobre a consciência com aqueles que não a mencionam. Se as premissas dizem respeito a fatos físicos ou neurais, então serão necessários alguns princípios de ponte ou ligações que conectem esses fatos com fatos sobre a consciência (Kim 1998). Ligações brutas, sejam elas nômicas ou meramente correlações bem confirmadas, poderiam fornecer uma ponte logicamente suficiente para inferir conclusões sobre a consciência. Mas provavelmente não nos permitiriam ver como ou por que tais conexões se dão, e assim estariam aquém de explicar completamente como a consciência existe (Levine 1983, 1993, McGinn 1991).

Poderíamos legitimamente pedir por mais, em particular por alguma abordagem que tornasse inteligível por que essas ligações se dão, e talvez por que não poderiam deixar de ser assim. Um modelo em dois estágios familiar para explicar as macropropriedades em termos de microsubstratos é invocado frequentemente. No primeiro passo, analisa-se a macropropriedade em termos de condições funcionais, e então, no segundo estágio, mostra-se que as microestruturas obedecendo as leis de seu próprio nível é nomicamente suficiente para garantir a satisfação das condições funcionais relevantes (Armstrong 1968, Lewis 1972).

As micropropriedades de coleções de moléculas de H₂O a 20°C são suficientes para satisfazer as condições para a liquidez da água que compõem. Além do mais, o modelo torna inteligível como a liquidez é produzida pelas micropropriedades. Poderia parecer que uma explicação satisfatória de como a consciência é produzida exigiria uma história similar em dois estágios. Sem isso, mesmo a dedutibilidade *a priori* poderia parecer explicativamente menos do que suficiente, embora a necessidade de uma tal história permaneça uma questão controversa (Block e Stalnaker 1999, Chalmers e Jackson 2001).

5.2 A lacuna explicativa

Seguindo Joseph Levine (1983), nossa incapacidade atual de suprir uma ligação inteligível adequada é por vezes descrita como a existência de uma *lacuna explicativa*, e como indicando nosso entendimento incompleto de como a consciência poderia depender de um substrato não-consciente, especialmente um substrato físico. A afirmação básica de que existe uma lacuna admite muitas variações de generalidade e, portanto, de força.

Talvez em sua forma mais fraca, ela asseve um limite *prático* em nossas capacidades explicativas *atuais*; dadas as nossas teorias e modelos atuais, não podemos articular agora uma ligação inteligível. Uma versão mais forte faz uma afirmação *em princípio* sobre as nossas *capacidades humanas*, asserindo assim que, dados os nossos limites cognitivos humanos, nunca seremos capazes de ultrapassar a lacuna. Para nós, ou para criaturas cognitivamente semelhantes a nós, deve permanecer um mistério residual (McGinn 1991). Colin McGinn (1995) defendeu que, dada a natureza inerentemente espacial, tanto de nossos conceitos perceptuais quanto dos conceitos científicos que deles derivamos, nós, humanos, não somos conceitualmente ajustados para entender a natureza da ligação psicofísica. Fatos sobre essa ligação estão tão cognitivamente fechados a nós quanto os fatos sobre multiplicação ou sobre raízes quadradas estão para os tatus. Eles não se enquadram em nosso repertório conceitual e cognitivo. Uma versão ainda mais forte da asserção da lacuna remove a restrição à nossa natureza cognitiva e nega *em princípio* que a lacuna possa ser preenchida por *qualquer agente cognitivo*.

Aqueles que defendem a existência da lacuna discordam entre si sobre que conclusões metafísicas, se alguma, seguem-se dos nossos supostos limites epistêmicos. O próprio Levine relutou em extrair quaisquer conclusões ontológicas antifisicalistas (Levine 1993, 2001). Por outro lado, alguns neodualistas têm tentado usar a existência da lacuna para refutar o fisicalismo (Foster 1996, Chalmers 1996). Quanto mais forte a premissa epistêmica, maior a esperança de se derivar uma conclusão metafísica. Assim, não surpreendentemente,

conclusões dualistas são frequentemente sustentadas por apelos à suposta impossibilidade *em princípio* de se fechar a lacuna.

Se pudéssemos ver, em bases *a priori*, que não há como a consciência ser inteligivelmente explicada como surgindo do físico, então não seria um grande passo concluir que isso não pode ser o caso (Chalmers 1996). Entretanto, a própria força dessa alegação epistêmica torna difícil chegar, sem incorrer em petição de princípio, ao resultado metafísico. Assim, aqueles que desejam usar a alegação forte da existência de uma lacuna *em princípio* para refutar o fisicalismo têm de encontrar bases independentes para sustentá-la. Alguns apelaram para argumentos de conceitabilidade [*conceivability*] em suporte, tal como a presumida conceitabilidade de zumbis molecularmente idênticos a humanos conscientes, mas desprovidos de toda consciência fenomenal (Campbell 1970, Kirk 1974, Chalmers 1996). Outros argumentos em suporte invocam a suposta natureza não-funcional da consciência, e daí sua alegada resistência ao método científico padrão de explicar propriedades complexas (por exemplo, dominância genética) em termos de condições funcionais realizadas fisicamente (Block 1980a, Chalmers 1996). Esses argumentos antifisicalistas evitam incorrer em petição de princípio, mas repousam, eles próprios, em alegações e intuições que são controversas e não completamente independentes do ponto de vista básico que se tem sobre o fisicalismo. Discussões sobre o tópico continuam ativas e em andamento.

Nossa incapacidade atual de ver qualquer maneira de preencher a lacuna pode exercer alguma atração em nossas intuições, mas pode simplesmente refletir os limites de nossa teorização atual, antes que uma barreira intransponível em princípio (Dennett 1991). Além do mais, alguns fisicalistas defenderam que lacunas explicativas devem ser esperadas, e são até implicadas, por versões plausíveis do fisicalismo ontológico, versões essas que tratam os agentes humanos como sistemas cognitivos fisicamente realizados, com limites inerentes que derivam de sua origem evolutiva e do seu modo contextual e situado de entendimento (Van Gulick 1985, 2003; McGinn 1991, Papineau 1995, 2002). Nessa perspectiva, antes de refutar o fisicalismo, a existência de lacunas explicativas pode confirmá-lo. A discussão e o desacordo nesses tópicos permanecem ativos e em andamento.

5.3 Explicação reducionista e não-reducionista

Como a necessidade de uma ligação inteligível já mostrou, uma dedutibilidade *a priori* não é, por si só, claramente suficiente para uma explicação bem sucedida (Kim 1980), nem é claramente necessária. Alguma ligação lógica mais fraca poderia ser suficiente em muitos contextos explicativos. Por vezes, contando o suficiente de uma história sobre como os fatos de um tipo dependem dos de outro podemos ficar satisfeitos de que os últimos de fato causam ou realizam os primeiros, mesmo que não possamos estritamente deduzir completamente os primeiros dos últimos.

A dedução interteórica estrita foi tomada como norma redutiva pela abordagem empirista lógica da unidade da ciência (Putnam e Oppenheim 1958), mas nas décadas mais recentes uma imagem não-reducionista mais frouxa das relações entre as várias ciências tem ganhado preferência. Em particular, materialistas não-reducionistas têm defendido a assim chamada “autonomia das ciências especiais” (Fodor 1974) e a ideia de que entender o mundo natural exige que usemos uma diversidade de sistemas conceituais e representacionais, que podem não ser estritamente intertraduzíveis ou capazes de serem postos na correspondência estrita exigida pelo antigo paradigma dedutivo das relações entre níveis (Putnam 1975).

A economia é frequentemente citada como um exemplo (Fodor 1974, Searle 1992). Fatos econômicos podem ser realizados por processos físicos subjacentes, mas ninguém pede seriamente que sejamos capazes de deduzir os fatos econômicos relevantes das descrições detalhadas de suas bases físicas subjacentes, ou que sejamos capazes de colocar os conceitos e o vocabulário da economia em correspondência estrita com aqueles das ciências físicas.

Apesar disso, nossa incapacidade dedutiva não é vista como motivo para suspeitas ontológicas; não há um problema da relação “dinheiro-matéria”. Tudo o que exigimos é algum entendimento geral e menos que dedutivo de como propriedades e relações econômicas poderiam ter por base propriedades e relações físicas. Assim, poderíamos optar por um critério similar para interpretar a questão de como a consciência pode existir e o que conta como explicar como a consciência poderia ser causada ou realizada por itens não-conscientes.

No entanto, alguns críticos, tais como Kim (1987), têm contestado a coerência de qualquer concepção que pretenda ser tanto não-redutiva quanto fisicalista, embora defensores dessas concepções tenham por sua vez replicado (Van Gulick 1993).

Outros têm defendido que a consciência é especialmente resistente à explicação em termos físicos, devido às diferenças inerentes entre nossos modos de compreensão subjetivos e objetivos. Thomas Nagel (1974) ficou famoso ao defender que existem limites inevitáveis impostos à nossa capacidade de entender a fenomenologia da experiência do morcego devido à nossa incapacidade de tomar empaticamente uma perspectiva experiencial como aquela que caracteriza a experiência auditiva ecolocalizadora que o morcego tem de seu mundo. Dada nossa incapacidade de submeter-nos a uma experiência similar, podemos ter no máximo um entendimento parcial da natureza de tal experiência. Nenhuma quantidade de conhecimento colhida pela perspectiva objetiva externa de terceira pessoa das ciências naturais será supostamente suficiente para nos permitir entender o que o morcego pode entender de sua própria experiência a partir do seu ponto de vista subjetivo interno de primeira pessoa.

5.4 Expectativas de sucesso explicativo

A questão de como a consciência pode existir subdivide-se, assim, em uma família diversificada de questões mais específicas, dependendo do tipo ou característica específica da consciência que se busca explicar, das restrições específicas que se faz quanto ao escopo das explicações [*explanans*] e ao critério que se usa para definir o sucesso explicativo. Algumas das variantes resultantes parecem mais fáceis de responder que outras. O progresso pode parecer provável em alguns dos assim chamados “problemas fáceis” da consciência, tais como explicar a dinâmica da consciência de acesso em termos da organização funcional ou computacional do cérebro (Baars 1988). Outros podem parecer menos tratáveis, especialmente o assim chamado “problema difícil” (Chalmers 1995), que é mais ou menos o de dar uma abordagem inteligível que nos permita ver de uma maneira intuitivamente

satisfatória como a consciência fenomenal ou “como é” pode surgir de processos físicos ou neurais no cérebro.

Respostas positivas a algumas versões das questões de como a consciência pode existir parecem ao alcance da mão, mas outras parecem permanecer profundamente frustrantes. Nem deveríamos supor que toda versão tem uma resposta positiva. Se o dualismo é verdadeiro, então, ao menos em alguns de seus tipos, a consciência poderia ser básica e fundamental. Se assim o for, não seremos capazes de explicar como ela surge de itens não-conscientes, dado que ela simplesmente não surge dessa forma.

A concepção que adotamos sobre as expectativas de se explicar a consciência dependerá tipicamente da perspectiva adotada. Fisicalistas otimistas provavelmente verão os lapsos explicativos atuais como mero reflexo de um estágio inicial de investigação, a serem certamente remediados num futuro não tão distante (Dennett 1991, Searle 1992, P. M. Churchland 1995). Para os dualistas, aqueles mesmos impasses significarão a falência do programa fisicalista e a necessidade de reconhecer a consciência como por si só um constituinte fundamental da realidade (Robinson 1982, Foster 1989, 1996, Chalmers 1996). O que se vê depende em parte de onde se está, e o projeto em andamento de explicar a consciência será acompanhado por contínuo debate sobre o seu estatuto e as expectativas de sucesso.

6. A questão funcional: *Por que a consciência existe?*

A questão funcional, ou do *porquê*, pergunta sobre o *valor* ou *papel* da consciência, e assim pergunta indiretamente sobre a sua origem. Ela tem uma *função* e, se tem, qual é? Ela faz alguma diferença na operação dos sistemas onde está presente e, se faz, por que e como faz? Se a consciência existe como uma característica complexa de sistemas biológicos, então seu valor adaptativo é provavelmente relevante na explicação de sua origem evolutiva, ainda que, evidentemente, sua função atual, se tem alguma, não precisa ser a mesma que ela talvez teve quando surgiu. Funções adaptativas frequentemente mudam no decorrer do tempo

biológico. Questões sobre o valor da consciência também têm uma dimensão *moral* em ao menos dois sentidos. Somos inclinados a ver o estatuto moral de um organismo como ao menos parcialmente determinado pela natureza e extensão em que é consciente, e estados conscientes, principalmente estados conscientes afetivos, tais como prazeres e dores, desempenham um papel central em muitas das abordagens do valor que subjazem a teoria moral (Singer 1975).

Assim como as questões do que é e como pode existir, a pergunta de por que a consciência existe também coloca um problema geral que se subdivide em uma variedade de indagações mais específicas. Na medida em que as várias formas de consciência, tais como consciência de acesso, fenomenal ou metamental, são distintas e separáveis – o que permanece uma questão aberta – elas provavelmente também diferem em seus papéis e valores específicos. Portanto, a questão de por que a consciência existe pode muito bem não ter uma resposta única ou uniforme.

6.1 Estatuto causal da consciência

Talvez a questão mais básica colocada por qualquer versão da pergunta de por que existe é se a consciência do tipo relevante tem ou não algum impacto causal. Se ela não tem nenhum efeito e não faz nenhuma diferença causal, seja ela qual for, então ela pareceria incapaz de desempenhar qualquer papel significativo nos sistemas e organismos onde está presente, eliminando assim, de partida, a maior parte das questões relativas ao seu possível valor. A ameaça de irrelevância epifenomenal também não poderia, dessa forma, ser simplesmente descartada como uma não-opção óbvia, uma vez que pelo menos algumas formas de consciência são seriamente defendidas na literatura recente como não tendo nenhum caráter causal. (Ver o verbete sobre “epifenomenalismo”, por Robinson 2011.) Esses problemas são levantados principalmente em relação aos qualia e à consciência qualitativa (Huxley 1874, Jackson 1982, Chalmers 1996), mas desafios também são apresentados ao estatuto causal das outras, incluindo a consciência metamental (Velmans 1991).

Argumentos tanto metafísicos quanto empíricos são oferecidos para sustentar ambas as posições. Entre os primeiros, estão aqueles que apelam a intuições sobre a conceitabilidade e a possibilidade lógica de zumbis, ou seja, de seres cujo comportamento, organização funcional e estrutura física até o nível molecular são idênticos a agentes humanos normais, mas que não possuem nenhum qualia ou consciência qualitativa. Alguns (Kirk 1970, Chalmers 1996) afirmam que tais seres são possíveis em mundos que compartilham todas as nossas leis físicas, mas outros negam isso (Dennett 1991, Levine 2001). Se são possíveis em tais mundos, então parece seguir-se que, mesmo no nosso mundo, os qualia não afetam o curso dos eventos físicos, incluindo os que constituem nossos comportamentos humanos. Se esses eventos se desenrolam da mesma forma, estejam ou não presentes os qualia, então os qualia parecem ser inertes ou epifenomenais, ao menos em relação aos eventos do mundo físico. Porém, esses argumentos e as intuições sobre zumbis nas quais se assentam são controversos, e a correção [soundness] deles continua em disputa (Searle 1992, Yablo 1998, Balog 1999).

Argumentos de um tipo bem mais empírico têm contestado o estatuto causal da consciência metamental, ao menos na medida em que sua presença pode ser medida pela habilidade de descrever o estado mental de alguém. Argumenta-se que evidências científicas mostram que uma consciência desse tipo não é nem necessária para nenhum tipo de habilidade mental nem ocorre suficientemente cedo para servir de causa dos atos ou processos tipicamente tomados como seus efeitos (Velmans 1991). De acordo com aqueles que defendem esses argumentos, os tipos de habilidades mentais que normalmente são vistas como necessitando a consciência podem todas ser realizadas inconscientemente sem a autoconsciência supostamente necessária.

Ademais, mesmo quando a consciência autorreflexiva está presente, ela supostamente ocorre tarde demais para ser a causa das ações relevantes, ao invés de seu resultado ou, no máximo, um efeito conjunto de alguma causa anterior compartilhada (Libet 1985). A autoconsciência ou consciência metamental, de acordo com esses argumentos, mostra-se um efeito psicológico posterior ao invés de uma causa iniciadora, mais como uma impressão *post*

facto ou o resultado exibido na tela do computador do que as efetivas operações do processador que produziram tanto a resposta do computador quanto o que está na tela.

Mais uma vez, os argumentos são controversos, e tanto os supostos dados quanto sua interpretação estão sujeitos a intensos desacordos (ver Flanagan 1992, e comentários acompanhando Velmans 1991). Apesar de os argumentos empíricos, assim como o dos zumbis, exigirem que consideremos seriamente se algumas formas de consciência são ou não causalmente menos potentes do que normalmente se presume, muitos teóricos veem os dados empíricos como não sendo uma ameaça real ao caráter causal da consciência.

Se os epifenomenalistas estão errados e a consciência, em suas várias formas, é de fato causal, então que tipos de efeitos ela tem e que diferença faz? Como os processos mentais que envolvem o tipo relevante de consciência diferem daqueles que não o envolvem? Quais funções a consciência poderia desempenhar? As próximas seis seções (6.2 – 6.7) discutem algumas das respostas mais comuns. Apesar das várias funções se sobreporem até certo ponto, cada uma delas é distinta, e elas diferem também quanto aos tipos de consciência com os quais cada uma está mais adequadamente ligada.

6.2 Controle flexível

Flexibilidade e sofisticação de controle aumentadas. Processos mentais conscientes parecem fornecer formas de controle altamente flexíveis e adaptativas. Ainda que processos automáticos inconscientes possam ser extremamente eficientes e rápidos, eles normalmente operam de maneira mais fixa e predeterminada do que aqueles que envolvem autoconsciência (Anderson 1983). A consciência é, portanto, da maior importância quando estamos lidando com situações novas e problemas ou demandas não previamente encontrados (Penfield 1975, Armstrong 1981).

As explicações convencionais da aquisição de habilidades ressaltam a importância da consciência durante a fase inicial de aprendizado, que gradualmente dá lugar a processos

mais automáticos, do tipo que requer pouca atenção ou supervisão consciente (Schneider e Shiffrin 1977). O processamento consciente permite a construção ou compilação de rotinas especificamente preparadas a partir de unidades elementares, bem como o controle deliberado de suas execuções.

Há uma troca bem conhecida entre flexibilidade e velocidade; processos conscientes controlados adquirem sua versatilidade sob medida às custas de serem lentos e cansativos, em contraste com a rapidez fluida das operações mentais inconscientes automáticas (Anderson 1983). Os crescimentos relevantes na flexibilidade parecem mais intimamente conectados com o tipo metamental, ou de ordem superior [*higher-order*], de consciência, na medida em que a melhora da habilidade de controlar processos depende do aumento de autoconsciência. No entanto, flexibilidade e modos sofisticados de controle também podem ser associados às formas de consciência fenomenal e de acesso.

6.3 Coordenação social

Capacidade aumentada para coordenação social. A consciência de tipo metamental pode envolver não apenas um acréscimo de autoconsciência, mas também uma melhora na compreensão dos estados mentais de outras criaturas com mentes, principalmente de outros membros do grupo social (Humphreys 1982). Criaturas que são conscientes no sentido metamental relevante não apenas possuem crenças, motivações, percepções e intenções, mas também compreendem como é ter tais estados e são conscientes tanto delas mesmas como de outros como os possuindo.

Esse acréscimo de conhecimento mutuamente compartilhado da mente um do outro permite que os organismos relevantes interajam, cooperem e comuniquem-se de formas muito mais avançadas e adaptativas. Apesar de a consciência metamental ser o tipo mais obviamente associado a essa função socialmente coordenativa, a consciência narrativa, do tipo associado ao fluxo de consciência, também é certamente relevante, na medida em que

envolve a aplicação na própria criatura de habilidades interpretativas que surgem em parte de suas aplicações sociais (Ryle 1949, Dennett 1978, 1992).

6.4 Representação integrada

Representação mais unificada e densamente integrada da realidade. A experiência consciente nos apresenta um mundo de objetos existindo independentemente no espaço e no tempo. Esses objetos nos são apresentados tipicamente de uma forma multimodal, que envolve a integração de informações de vários canais sensoriais, assim como do conhecimento de base e da memória. A experiência consciente não nos apresenta propriedades ou características isoladas, mas sim objetos e eventos situados num mundo independente em fluxo, e ela faz isso incorporando em sua organização e dinâmica experiencial a densa rede de relações e interconexões que coletivamente constituem a estrutura significativa de um mundo de objetos (Kant 1787, Husserl 1913, Campbell 1997).

Evidentemente, nem toda informação sensorial precisa ser experienciada para ter um efeito adaptativo no comportamento. Conexões adaptativas sensorio-motoras não-experienciais podem ser encontradas tanto em organismos simples quanto em alguns dos processos mais diretos e reflexivos de organismos superiores. Mas quando há experiência, ela fornece uma representação mais unificada e integrada da realidade, que normalmente propicia vias mais abertas de resposta (Lorenz 1977). Considere, por exemplo, a representação do espaço em um organismo cujos canais de entrada sensorial são simplesmente ligados ao movimento ou à orientação de poucos mecanismos fixos, tais como aqueles para alimentação ou para agarrar presas, e compare isso com um organismo capaz de usar sua informação espacial para a navegação flexível do seu ambiente e para quaisquer outros objetivos ou metas espacialmente relevantes que possa ter, como quando uma pessoa examina visualmente seu escritório ou sua cozinha (Gallistel 1990).

É a representação desse último tipo que normalmente é disponibilizada pelo modo integrado de apresentação associado à experiência consciente. A unidade do espaço

experienciado é apenas um exemplo do tipo de integração associado à nossa consciência de um mundo objetivo. (Ver o verbete sobre a “unidade da consciência”, por Brook e Raymont 2010.)

Esse papel ou valor integrativo é mais obviamente associado com o tipo fenomenal de consciência, mas, na medida em que os qualia desempenham um papel na apresentação de objetos na experiência, esse papel unificador também é associado à consciência nos sentidos qualitativo e de “como-é” [*what it's like*]. Ademais, ele está intimamente ligado à transparência da experiência, principalmente a transparência semântica, descrita em resposta à questão sobre o que é a consciência (Van Gulick 1993).

6.5 Acesso informacional

Mais acesso informacional global. A informação carregada por estados mentais conscientes normalmente está disponível para o uso de diversos subsistemas mentais e para a aplicação em um amplo rol de situações e ações potenciais (Baars 1988). A informação não-consciente é mais passível de ser encapsulada no interior de módulos mentais particulares e ficar disponível apenas para o uso concernente às aplicações diretamente conectadas às operações desse subsistema (Fodor 1983). Tornar uma informação consciente normalmente amplia a esfera de sua influência e o número de formas em que pode ser usada para guiar e conformar adaptativamente tanto o comportamento interno quanto externo. O caráter consciente de um estado pode ser em parte uma questão daquilo que Dennett chamou de “celebridade cerebral”, ou seja, a capacidade desse estado de ter um impacto compatível com o conteúdo de outros estados mentais.

Essa função particular está ligada mais diretamente, e por definição, à noção de consciência de acesso (Block 1995), mas a consciência metamental, bem como as formas fenomenal e qualitativa, parecem todas associadas a esse crescimento na disponibilidade de informação (Armstrong 1981, Tye 1985).

6.6 Livre-arbítrio

Aumento da liberdade de escolha ou livre-arbítrio. A questão do livre-arbítrio permanece um problema filosófico perene, não apenas no que diz respeito à sua existência ou não, mas até mesmo no que se refere a em que isso poderia ou deveria consistir (Dennett 1984, van Inwagen 1983, Hasker 1999, Wegner 2002). (Ver o verbete sobre “Livre-arbítrio”, por O’Connor 2011.) Talvez a própria noção de livre-arbítrio continue muito obscura e controversa para esclarecer o que quer que seja sobre o papel da consciência, mas há uma intuição tradicional de que as duas coisas estão intimamente ligadas.

A consciência é vista como abrindo um reino de possibilidades, uma esfera de opções na qual o eu consciente pode escolher ou agir livremente. A consciência parece, no mínimo, uma pré-condição necessária para qualquer forma de liberdade ou autodeterminação desse tipo (Hasker 1999). Como alguém poderia exercer o tipo exigido de livre-arbítrio e, ao mesmo tempo, permanecer apenas no interior do domínio inconsciente? Como alguém poderia determinar sua própria vontade sem ser consciente dela e das opções disponíveis para moldá-la?

A liberdade para escolher a ação a tomar e a habilidade de determinar a própria natureza e desenvolvimento futuro podem admitir diversos graus e variações interessantes, ao invés de serem simplesmente uma questão de tudo ou nada, e as várias formas ou níveis de consciência podem estar associadas a graus e tipos correspondentes de liberdade e autodeterminação (Dennett 1984, 2003). A ligação com a liberdade parece mais forte para a forma metamental de consciência, dada a sua ênfase na autoconsciência, mas conexões potenciais também parecem igualmente possíveis para a maior parte das outras formas.

6.7 Motivação intrínseca

Estados intrinsecamente motivacionais. Ao menos alguns estados conscientes parecem ter intrinsecamente a força motora que têm. Em particular, os papéis funcionais e

motivacionais de estados afetivos conscientes, tais como prazeres e dores, parecem intrínsecos ao seu caráter experiencial e inseparáveis de suas propriedades qualitativas e fenomenais, ainda que isso tenha sido colocado em questão (Nelkin 1989, Rosenthal 1991). O aspecto motivacional positivamente atrativo do prazer parece uma parte de sua sensação fenomenal diretamente experienciada, assim como o caráter afetivo negativo da dor, ao menos no caso da experiência não-patológica normal.

Há considerável controvérsia sobre o quanto o sentimento e a força motora da dor podem ser dissociados em casos anormais, e alguns negam a própria existência de tais aspectos intrinsecamente motivacionais (Dennett 1991). No entanto, ao menos no caso normal, a força motivacional negativa da dor *parece* constitutiva da sensação da própria experiência.

Exatamente *como* isso poderia ser o caso continua não sendo muito claro, e talvez a aparente força motivacional intrínseca e diretamente experienciada seja ilusória. Mas se for real, então pode ser um dos sentidos mais importantes e evolutivamente mais antigos em que a consciência faz diferença para os sistemas e processos mentais onde está presente (Humphreys 1992).

Outras sugestões foram feitas sobre as possíveis funções e valor da consciência, e essas seis evidentemente não exaurem as opções. No entanto, elas estão entre as hipóteses recentes mais proeminentes, e fornecem um levantamento satisfatório dos tipos de respostas que têm sido oferecidas à questão de por que a consciência existe por aqueles que acreditam que a consciência de fato faz alguma diferença.

6.8 Papéis constitutivos e contingentes

Um ponto adicional exige esclarecimento sobre os vários sentidos em que as funções propostas poderiam responder à pergunta de por que a consciência existe. Em particular, deve-se distinguir entre casos *constitutivos* e casos de *realização contingente*. No primeiro, desempenhar o papel constitui ser consciente no sentido relevante, enquanto que no segundo

a consciência de um dado tipo é apenas uma entre várias formas em que o papel exigido poderia ser realizado (Van Gulick 1993).

Por exemplo: tornar a informação disponível globalmente para uso de uma ampla variedade de subsistemas e aplicações comportamentais pode constituir seu caráter consciente no sentido de acesso. Por outro lado, mesmo que as formas qualitativas e fenomenais de consciência envolvam uma representação altamente unificada e densamente integrada da realidade objetiva, pode ser possível produzir representações que tenham essas características funcionais, mas que não sejam de natureza qualitativa ou fenomenal.

O fato de que *em nós* os modos de representação com essas características também tenham propriedades qualitativas e fenomenais pode refletir fatos históricos contingentes sobre a solução particular de modelo que calhou surgir em nossa linhagem evolutiva. Se isso for verdade, deve haver meios bem diversos de se atingir um resultado comparável sem a consciência qualitativa ou fenomenal. Não é claro que essa seja a forma adequada de se pensar a consciência qualitativa e fenomenal; talvez a associação com representações unificadas e densamente integradas seja de fato tão íntima e constitutiva quanto parece ser o caso na consciência de acesso (Carruthers 2000). Independentemente de como essa questão seja resolvida, é importante não misturar abordagens constitutivas com abordagens de realização contingente quando se investiga a função da consciência e se responde à questão de por que ela existe (Chalmers 1996).

7. Teorias da consciência

Em resposta às perguntas sobre o que é, como é e por que existe, muitas teorias da consciência foram propostas nos últimos anos. Entretanto, nem todas as teorias da consciência são teorias da mesma coisa. Elas variam não apenas quanto aos tipos de consciência que têm como seu objeto, mas também em suas finalidades teóricas.

Talvez a maior divisão seja entre teorias metafísicas gerais, que pretendem situar a consciência no interior de um esquema ontológico abrangente da realidade, e teorias

específicas, que oferecem explicações detalhadas de sua natureza, características e função. A linha separando esses dois tipos de teoria é um tanto obscura, principalmente na medida em que muitas teorias específicas carregam ao menos alguns comprometimentos com questões metafísicas mais gerais. No entanto, é útil manter a divisão em mente quando se faz um levantamento do leque de ofertas teóricas correntes.

8. Teorias metafísicas da consciência

Teorias metafísicas gerais oferecem respostas para a versão consciente do problema mente-corpo: qual o estatuto ontológico da consciência em relação ao mundo da realidade física? Em grande medida, as respostas disponíveis acompanham as alternativas convencionais ao problema mente-corpo, incluindo as principais versões de dualismo e fisicalismo.

8.1 Teorias dualistas

Teorias dualistas veem ao menos alguns aspectos da consciência como estando fora do reino físico, mas formas específicas de dualismo diferem a respeito de quais são, exatamente, esses aspectos. (Ver o verbete por Robinson 2011.)

O *dualismo de substância*, tal como o dualismo tradicional cartesiano (Descartes 1644), afirma a existência de ambas as substâncias, física e não-física. Essas teorias implicam a existência de mentes ou egos não-físicos como entidades às quais a consciência é inerente. Apesar do dualismo de substância estar em grande medida em desuso no momento, existem alguns defensores contemporâneos (Swinburne 1986, Foster 1989, 1996).

O *dualismo de propriedade*, em suas várias versões, goza de maior popularidade atualmente. Todas as teorias desse tipo afirmam a existência de propriedades conscientes que não são nem idênticas nem redutíveis a propriedades físicas, mas que podem, entretanto, ser instanciadas exatamente pelas mesmas coisas que instanciam propriedades físicas. Nesse

sentido, elas podem ser classificadas como teorias de *duplo aspecto*. Elas tomam algumas partes da realidade – organismos, cérebros, estados ou processos neurais – como instanciando propriedades de dois tipos distintos e desconexos: as físicas e as conscientes, fenomenais ou qualitativas. Teorias de duplo aspecto, ou dualistas de propriedades, podem ser de ao menos três tipos diferentes.

O *dualismo de propriedades fundamental* vê propriedades mentais conscientes como constituintes básicos da realidade, ao lado de propriedades físicas, como carga eletromagnética. Elas podem interagir de maneira causal e legiforme [*law-like*] com outras propriedades fundamentais, tais como as da física, mas ontologicamente a existência delas não é nem dependente nem derivada de nenhuma outra propriedade (Chalmers 1996).

O *dualismo de propriedades emergentes* vê as propriedades conscientes como surgindo de organizações complexas de constituintes físicos, mas de uma forma radical tal que o resultado emergente é algo para além de suas causas físicas, não previsível *a priori* e não explicável em termos de suas naturezas estritamente físicas. A coerência de tais abordagens emergentistas foi desafiada (Kim 1998), mas elas têm defensores (Hasker 1999).

O *dualismo de propriedades monista e neutro* trata tanto propriedades mentais conscientes quanto propriedades físicas como, de alguma forma, dependentes ou derivadas de um nível mais básico de realidade, que não é, ele mesmo, nem mental nem físico (Russell 1927, Strawson 1994). Entretanto, se entendemos o dualismo como a tese de que há dois reinos distintos de entidades ou propriedades fundamentais, então o monismo neutro talvez não devesse ser classificado como uma versão de dualismo de propriedades, uma vez que não vê nem as propriedades mentais nem as físicas como últimas ou fundamentais.

O *pampsiquismo* pode ser visto como um quarto tipo de dualismo de propriedades, uma vez que toma todos os constituintes da realidade como tendo algumas propriedades psíquicas, ou ao menos protopsíquicas, distintas de qualquer propriedade física que possam ter (Nagel 1979). De fato, o monismo neutro poderia ser consistentemente combinado com alguma versão de *pamprotopsiquismo* (Chalmers 1996), na qual os aspectos protometais dos micro-constituintes poderiam dar origem, em condições de combinação propícias, à

consciência em sentido pleno. (Ver o verbete sobre “pampsiquismo”, por Seager e Allen-Hermanson 2012.)

A natureza do aspecto protopsíquico relevante continua obscura e essas teorias enfrentam um dilema, se oferecidas na esperança de responder ao Problema Difícil [*Hard Problem*]. Ou as propriedades protopsíquicas envolvem o tipo de sensação qualitativa fenomenal que gera o Problema Difícil, ou elas não envolvem. Caso envolvam, é difícil entender como elas podem possivelmente ocorrer como propriedades ubíquas da realidade. Como pode um elétron ou um quark ter qualquer sensação experiencial desse tipo? No entanto, se as propriedades protopsíquicas não envolvem sensações desse tipo, então não é claro como seriam, de alguma forma, mais aptas do que as propriedades físicas para dar conta da consciência qualitativa ao solucionar o Problema Difícil.

Muitos argumentos foram apresentados a favor da teoria dualista e de outras teorias antifisicalistas da consciência. Alguns são, em grande parte, *a priori* em natureza, como aqueles que apelam para a suposta conceitabilidade de zumbis (Kirk 1970, Chalmers 1996) ou versões do argumento do conhecimento (Jackson 1982, 1986), que visa alcançar uma conclusão antifisicalista sobre a ontologia da consciência a partir dos aparentes limites em nossa habilidade de compreender inteiramente os aspectos qualitativos da experiência consciente através de abordagens físicas de terceira-pessoa dos processos cerebrais. (Ver Jackson 1998, 2004 para uma visão contrária; ver também os verbetes sobre “zumbis”, por Kirk 2012, e “Qualia: o argumento do conhecimento”, por Nida-Rümelin 2010.) Outros argumentos a favor do dualismo possuem bases mais empíricas, tais como aqueles que apelam à suposta lacuna causal nas cadeias de causação física no cérebro (Eccles e Popper 1977), ou aqueles que se baseiam em supostas anomalias na ordem temporal da experiência consciente (Libet 1982, 1985). Argumentos dualistas de ambos os tipos são muito questionados por fisicalistas (P. S. Churchland 1981, Dennett e Kinsbourne 1992).

8.2 Teorias fisicalistas

A maioria das outras teorias metafísicas da consciência é uma ou outra versão de algum tipo já conhecido de fisicalismo.

Teorias *eliminativistas* negam redutivamente a existência da consciência, ou ao menos a existência de alguns de seus tipos ou características comumente aceitos. (Ver o verbete sobre “o materialismo eliminativista”, por Ramsey 2011.) Os eliminativistas radicais rejeitam a própria noção de consciência como sendo confusa ou enganadora, e defendem que a distinção consciente/não-consciente não consegue cortar a realidade mental em suas articulações (Wilkes 1984, 1988). Eles entendem que a ideia de consciência é suficientemente equivocada para merecer ser eliminada e substituída por outros conceitos e distinções que melhor reflitam a verdadeira natureza da mente (P. S. Churchland 1983).

A maior parte dos eliminativistas é mais sofisticada em suas avaliações críticas. Ao invés de rejeitarem a noção completamente, questionam apenas algumas das características proeminentes que ela é comumente vista como envolvendo, tais como os qualia (Dennett 1990, Carruthers 2000), o eu consciente (Dennett 1992), ou o chamado “teatro cartesiano”, onde a sequência temporal da experiência consciente é projetada internamente (Dennett e Kinsbourne 1992). Portanto, eliminativistas mais modestos, como Dennett, normalmente combinam suas negações mais qualificadas com uma teoria positiva daqueles aspectos da consciência que tomam como reais, tal como acontece no Modelo de Múltiplas Camadas [*Multiple Drafts Model*] (seção 9.3 a seguir).

A *teoria da identidade*, ao menos a teoria da identidade psicofísica tipo-tipo [*type-type*] estrita, oferece uma outra opção fortemente redutivista ao identificar propriedades, estados e processos mentais conscientes com físicos, normalmente de natureza neural ou neurofisiológica. Se ter uma experiência consciente qualitativa do vermelho fenomenal é *simplesmente* estar em um estado cerebral com as propriedades neurofisiológicas relevantes, então essas propriedades experienciais são reais, mas a realidade delas é uma realidade diretamente física.

A *teoria da identidade tipo-tipo* é chamada assim porque identifica tipos ou propriedades mentais e físicas, da mesma forma que se identifica a propriedade de ser água com a propriedade de ser composto de moléculas de H₂O. Após um breve período de popularidade nos primeiros dias do fisicalismo contemporâneo, nos anos 1950 e 60 (Place 1956, Smart 1959), ela tem sido bem menos amplamente adotada devido a problemas como a objeção da realização múltipla, de acordo com a qual propriedades mentais são mais abstratas e por isso capazes de serem realizadas por estruturas de fundo ou substratos químicos diferentes (Fodor 1974, Hellman e Thompson 1975). Se uma e a mesma propriedade consciente pode ser realizada por diferentes propriedades neurofisiológicas (ou mesmo não-neurofisiológicas) em diferentes organismos, então as duas propriedades não podem ser estritamente idênticas.

Entretanto, a teoria da identidade tipo-tipo gozou um renascimento recente, ainda que modesto, ao menos no que se refere aos qualia, ou propriedades conscientes qualitativas. Isso se deu, em parte, porque tratar a ligação psico-física como uma identidade foi visto por alguns como propiciando uma forma de dissolver o problema da lacuna explicativa (Hill e McLaughlin 1998, Papineau 1995, 2003). Eles argumentam que se a propriedade qualitativa consciente e a propriedade neural são idênticas, então não há necessidade de explicar como a última causa ou faz surgir a primeira. Uma não *causa* a outra: uma *é* a outra. Dessa forma, não há lacuna a ser preenchida, e nenhuma explicação adicional é necessária. Identidades não são o tipo de coisa que pode ser explicada, uma vez que nada é idêntico a nada além de si mesmo, e não faz sentido perguntar porque algo é idêntico a si mesmo.

No entanto, outros contestam que o apelo à identidade tipo-tipo não elimina de forma tão óbvia a necessidade de uma explicação (Levine 2001). Mesmo quando duas descrições ou conceitos de fato se referem a uma e a mesma propriedade, pode-se ainda esperar, racionalmente, por alguma explicação dessa convergência, alguma explicação de como selecionaram uma e a mesma coisa, ainda que, inicialmente ou intuitivamente, não pareçam fazer isso. Em outros casos de identidades de propriedades empiricamente descobertas, tal como entre calor e energia cinética, há uma história a ser contada que explica a convergência

correferencial, e parece justo esperar o mesmo no caso psicofísico. Portanto, apelar para identidades tipo-tipo talvez não seja suficiente, por si só, para dissolver o problema da lacuna explicativa.

A maior parte das teorias fisicalistas da consciência não são nem eliminativistas nem baseadas em identidades tipo-tipo estritas. Elas reconhecem a realidade da consciência, mas procuram situá-la no interior do mundo físico, por meio de alguma relação psicofísica mais fraca que a estrita identidade de propriedade.

Entre as variantes comuns, estão aquelas que tomam a realidade da consciência como *superveniente* sobre o físico, como *composta do* físico ou como sendo *realizada pelo* físico.

Teorias *funcionalistas*, em particular, apoiam-se largamente na noção de *realização* para explicar a relação entre a consciência e o físico. De acordo com o funcionalismo, um estado ou processo conta como sendo de um dado tipo mental ou consciente em virtude do papel funcional que executa no interior de um sistema apropriadamente organizado (Block 1980a). Um dado estado físico realiza o tipo mental consciente relevante ao executar o papel apropriado no interior de um sistema físico mais amplo que o contém. (Ver o verbete sobre o “funcionalismo”, por Levin 2010.) O funcionalista apela muitas vezes a analogias com outras relações entre níveis, como entre o biológico e o bioquímico, ou entre o químico e o atômico. Em cada um desses casos, propriedades e fatos de um certo nível são realizados por interações complexas entre itens de um nível subjacente.

Os críticos do funcionalismo negam frequentemente que a consciência possa ser adequadamente explicada em termos funcionais (Block 1980a, 1980b, Levine 1983, Chalmers 1996). De acordo com alguns críticos, a consciência pode ter características funcionais interessantes, mas sua natureza não é essencialmente funcional. Essas afirmações são por vezes sustentadas por apelo à suposta possibilidade dos qualia ausentes ou invertidos, ou seja, a possibilidade de existirem seres funcionalmente equivalentes aos humanos normais, mas que possuem qualia invertidos ou não possuem nenhum qualia. O estatuto dessas possibilidades é controverso (Shoemaker 1981, Dennett 1990, Carruthers 2000), mas se elas

são aceitas, parecem suscitar um problema para o funcionalista. (Ver o verbete “Qualia”, por Tye 2009.)

Aqueles que apoiam o fisicalismo ontológico na relação de realização frequentemente combinam isso com uma visão não-redutivista, no nível conceitual ou representacional, que ressalta a autonomia das ciências especiais e os modos distintos de descrição e de acesso cognitivo que elas ensejam.

O *fisicalismo não-redutivista* desse tipo nega que os recursos teóricos e conceituais apropriados e adequados para lidar com fatos do nível de substrato subjacente, ou nível da realização, tenham de ser igualmente adequados para lidar com aqueles do nível realizado (Putnam 1975, Boyd 1980). Como foi visto anteriormente, em resposta à pergunta sobre como a consciência pode existir, pode-se acreditar que todos os fatos econômicos são fisicamente realizados sem, com isso, pensar que os recursos das ciências físicas oferecem todas as ferramentas cognitivas e conceituais de que precisamos para fazer economia (Fodor 1974).

O fisicalismo não-redutivista tem sido criticado pela suposta incapacidade de “pagar sua dívida com o fisicalismo” em moeda redutivista. Seu erro é supostamente o de não dar uma explicação adequada de como propriedades conscientes são, ou poderiam ser, realizadas por estruturas e processos neurais, físicos ou funcionais subjacentes (Kim 1987, 1998). Ele foi, na verdade, acusado de incoerência, devido à sua tentativa de combinar a afirmação da realização física com a negação da capacidade de explicitar essa relação de uma forma estrita e inteligível *a priori* (Jackson 2004).

Entretanto, como observado antes na discussão sobre a questão de como a consciência existe, os fisicalistas não-redutivistas reagem concordando que uma abordagem da realização psicofísica é de fato necessária, mas adicionam que a abordagem relevante pode diferir muito de uma dedutibilidade *a priori* e, ainda assim, ser suficiente para satisfazer nossas legítimas demandas explicativas (McGinn 1991, Van Gulick 1985). A questão continua em discussão.

9. Teorias específicas da consciência

Apesar de existirem várias teorias metafísicas/ontológicas gerais da consciência, a lista de teorias específicas detalhadas acerca de sua natureza é ainda mais longa e mais diversificada. Nenhuma investigação breve poderia chegar perto de ser exaustiva, mas seis tipos principais de teorias podem ajudar a indicar o leque básico de opções: teorias de ordem superior [*higher-order theories*], teorias representacionais, teorias cognitivas, teorias neurais, teorias quânticas e teorias não-físicas.

9.1 Teorias de ordem superior

Teorias de ordem superior (OS)¹ analisam a noção de estado mental consciente em termos de autoconsciência metamental reflexiva. A ideia central é que o que torna um estado mental M um estado mental consciente é o fato de estar acompanhado de um estado simultâneo e não-inferencial de ordem superior (isto é, metamental), cujo conteúdo é que se está no estado M no momento. Ter um desejo consciente por chocolate envolve ter dois estados mentais: é preciso ter tanto um desejo por chocolate, quanto um estado de ordem superior cujo conteúdo é que se está, no momento, tendo exatamente esse desejo. Estados mentais inconscientes são inconscientes precisamente por não termos os estados de ordem superior relevantes sobre eles. Seu caráter inconsciente consiste no fato de não estarmos reflexiva e diretamente cientes de estarmos em tais estados. (Ver o verbete sobre “teorias de ordem superior da consciência”, por Carruthers 2011.)

Teorias de ordem superior aparecem em duas variantes principais, que diferem quanto ao modo psicológico dos estados mentais constitutivos da consciência relevante. Teorias do pensamento de ordem superior (PenOS) tomam o estado de ordem superior requerido como sendo um metaestado assertórico similar ao pensamento (Rosenthal 1986, 1993). Teorias da percepção de ordem superior (PerOS) os veem mais como estados perceptivos,

¹ A sigla original é “HO”, de “*higher-order*”.

associados a uma espécie de sentido interno e a sistemas de monitoramento intramentais de algum tipo (Armstrong 1981, Lycan 1987, 1996).²

Cada uma tem suas relativas vantagens e problemas. Proponentes da PenOS ressaltam que não temos nenhum órgão de sentido interno, e afirmam que não experienciamos nenhuma qualidade sensível a não ser aquelas que nos são apresentadas pela percepção direcionada ao exterior. Proponentes da PerOS, por outro lado, podem argumentar que sua abordagem explica algumas das condições adicionais exigidas por teorias OS como sendo consequências naturais da natureza perceptiva dos estados de ordem superior relevantes. Em particular, a exigência de que o metaestado constitutivo da consciência seja não-inferencial e simultâneo ao objeto mental de nível inferior pode ser explicada pelas condições correspondentes que normalmente se aplicam à percepção. Percebemos o que está acontecendo agora, e o fazemos de uma forma que não envolve nenhuma inferência, ao menos não uma inferência explícita no nível pessoal. Essas condições são igualmente necessárias para a PenOS, mas não são explicadas por essa teoria, o que aparentemente dá uma vantagem explicativa ao modelo PerOS (Lycan 2004, Van Gulick 2000), ainda que alguns teóricos da PerOS argumentem contrariamente (Carruthers 2000).

Sejam quais forem seus respectivos méritos, tanto teorias PerOS quanto PenOS enfrentam alguns desafios comuns, incluindo o que poderia ser chamado de *problema da generalidade*. Ter um pensamento ou uma percepção de um dado item *X* – seja uma pedra, uma caneta ou uma batata – não faz, em geral, de *X* um *X* consciente. Ver uma batata em cima do balcão, ou pensar nela, não faz dela uma batata consciente. Por que deveria, então, o fato de ter um pensamento ou uma percepção de um determinado desejo ou memória tornar esse desejo ou essa memória conscientes? (Dretske 1995, Byrne 1997). Também não basta ressaltar que não utilizamos o termo “consciente” para pedras ou canetas que percebemos ou pensamos, mas apenas para estados mentais que percebemos ou a respeito dos quais pensamos (Lycan 1997, Rosenthal 1997). Isso pode bem ser verdade, mas o que falta é uma explicação de por que isso é adequado.

² As siglas originais são, respectivamente, “HOT” (*Higher-Order Thought*) e “HOP” (*Higher-Order Perception*).

A abordagem de ordem superior é mais claramente relevante para as formas metamentais de consciência, mas alguns de seus defensores acreditam que ela explica também outros tipos de consciência, incluindo os tipos mais subjetivos de consciência como-é e qualitativa. Uma estratégia comum é analisar os qualia como características mentais que são capazes de ocorrer inconscientemente; por exemplo, eles podem ser explicados como propriedades de estados internos, cujas relações estruturadas de semelhança fazem surgir crenças sobre semelhanças objetivas no mundo (Shoemaker 1975, 1990). Apesar dos qualia inconscientes poderem desempenhar o papel funcional, não precisa haver um como-é estar num estado que os possua (Nelkin 1989, Rosenthal 1991, 1997). De acordo com o teórico da OS, o caráter *como-é* [*what-it's-likeness*] surge apenas quando nos tornamos cientes de um estado de primeira ordem e de suas propriedades qualitativas através de um metaestado apropriado direcionado a ele.

Críticos da abordagem OS questionaram essa perspectiva, e alguns argumentaram que a noção de qualia inconsciente, na qual se baseia, é incoerente (Papineau 2002). Independentemente de quão satisfatória seja a abordagem dos qualia proposta pelas teorias OS, é importante ressaltar que a maior parte dos defensores da OS acreditam estar oferecendo uma teoria ampla da consciência, ou ao menos o cerne de uma teoria geral desse tipo, ao invés de simplesmente uma teoria restrita a alguns tipos metamentais especiais de consciência.

Outras variantes da teoria OS vão além das tradicionais versões PenOS e PerOS, incluindo aquelas que analisam a consciência em termos de pensamentos disposicionais de ordem superior, ao invés de pensamentos ocorrentes (Carruthers 2000). Outras apelam para um entendimento de ordem superior implícito, ao invés de explícito, e enfraquecem ou eliminam a pressuposição tradicional de que um metaestado precisa ser distinto e separado de seus objetos de ordem inferior (Gennaro 1995, Van Gulick 2000, 2004). Outras variantes da teoria OS continuam sendo oferecidas, e o debate entre defensores e críticos da abordagem básica continua vivo (Ver os artigos recentes em Gennaro 2004).

9.2 Teorias representacionalistas

Praticamente todas as teorias da consciência a veem como possuindo características representacionais, mas as assim chamadas teorias representacionalistas se definem pela afirmação mais forte de que suas características representacionais exaurem suas características mentais (Harman 1990, Tye 1995, 2000). De acordo com o representacionalista, estados mentais conscientes não possuem nenhuma propriedade mental além de suas propriedades representacionais. Portanto, dois estados conscientes ou experienciais que compartilham todas as suas propriedades representacionais não diferirão em nenhum aspecto mental.

A força exata dessa afirmação depende de como se interpreta a noção de ser “representacionalmente o mesmo”, havendo várias opções plausíveis de critérios. Pode-se defini-la, grosso modo, em termos de condições de satisfação ou de verdade, mas, assim entendida, a tese representacionalista parece claramente falsa. Existem várias maneiras em que estados podem compartilhar suas condições de satisfação ou verdade e, ainda assim, diferirem mentalmente, incluindo aquelas referentes aos modos de conceitualizar ou apresentar essas condições.

No extremo oposto, pode-se tomar dois estados como representacionalmente distintos se diferirem em qualquer característica que desempenhe um papel na operação ou função representacional deles. Numa interpretação liberal desse tipo, qualquer diferença nos portadores do conteúdo conta como diferença representacional, ainda que tenham o mesmo conteúdo intencional ou representacional, podem diferir apenas em seus *meios* ou *modos* de representação, e não em seus *conteúdos*.

Evidentemente, essa interpretação aumenta a plausibilidade da afirmação de que as propriedades representacionais de um estado consciente exaurem suas propriedades mentais, mas ao custo de enfraquecer significativamente, ou mesmo trivializar, a tese. Portanto, o representacionalista parece precisar de uma interpretação da *igualdade representacional* que vá além das meras condições de satisfação, e que reflita todos os

aspectos intencionais e relativos ao conteúdo da representação, sem ser sensível a meras diferenças em características subjacentes sem conteúdo dos processos no nível de realização. Portanto, a maior parte dos representacionistas estipulam condições para experiências conscientes que incluem, ao lado de uma condição relativa ao conteúdo, algum papel causal, ou exigências de configuração adicionais (Tye 1995, Dretske 1995, Carruthers 2000). Outros representacionistas aceitam a existência dos qualia, mas tratam-nos como propriedades objetivas que objetos externos são representados como possuindo. Ou seja, tratam-nos como *propriedades representadas*, ao invés de *propriedades de representações* ou de estados mentais (Dretske 1995, Lycan 1996).

O representacionismo pode ser visto como uma forma qualificada de eliminativismo, uma vez que nega a existência de propriedades de um tipo que estados mentais conscientes são usualmente pensados como possuindo – ou ao menos como parecendo possuir – a saber, aquelas que são mentais, mas não representacionais. Qualia, ao menos se entendidos como propriedades monádicas intrínsecas de estados conscientes acessíveis à introspecção, pareceriam ser os alvos mais óbvios dessa eliminação. De fato, parte da motivação do representacionismo é mostrar que podemos acomodar todos os fatos sobre a consciência, talvez dentro de um quadro fisicalista, sem precisarmos encontrar um lugar para os qualia ou para qualquer outra suposta propriedade mental não-representacional (Dennett 1990, Lycan 1996, Carruthers 2000).

O representacionismo tem sido muito popular nos últimos anos e possui muitos defensores, mas continua altamente controverso e contraintuitivo quanto a casos e experimentos mentais cruciais (Block 1996). Em particular, a possibilidade de qualia invertidos fornece um caso crucial para teste. Para antirrepresentacionistas, a mera possibilidade lógica de qualia invertidos mostra que estados conscientes podem diferir em um aspecto mental significativo e, ao mesmo tempo, coincidir representacionalmente. Representacionistas, em resposta, ou negam a possibilidade de tal inversão, ou sua alegada significação (Dretske 1995, Tye 2000).

Muitos outros argumentos foram levantados a favor ou contra o representacionalismo, tais como aqueles a respeito das percepções em diferentes modalidades sensoriais do mesmo estado de coisas – ver e tocar o mesmo cubo –, o que parece envolver diferenças mentais distintas de como os estados relevantes representam o mundo como sendo (Peacocke 1983, Tye 2003). Em cada caso, ambos os lados podem reunir fortes intuições e engenho argumentativo. Um vívido debate continua.

9.3 Teorias cognitivas

Modelos que procuram explicar a consciência em termos de processos cognitivos têm sido oferecidos tanto por filósofos quanto por psicólogos.

O exemplo filosófico de maior destaque é o Modelo de Múltiplas Camadas (MMC) da consciência, proposto por Daniel Dennett (1991).³ Ele combina elementos tanto do representacionalismo, quanto da teoria de ordem superior, mas o faz de tal maneira que o distancia, de forma interessante, das versões convencionais de ambos. O MMC inclui muitas características distintas, mas inter-relacionadas.

Seu nome reflete o fato de que, num momento qualquer, determinações de conteúdos de vários tipos ocorrem em todo o cérebro. O que faz alguns desses conteúdos serem conscientes não é o fato de ocorrerem num local espacial ou funcional privilegiado – o chamado “Teatro Cartesiano” –, nem o fato de ocorrerem de um modo ou num formato especial: tudo isso é negado pelo MMC. Ao contrário, isso é uma questão do que Dennett chama de “celebridade cerebral” [*cerebral celebrity*], isto é, o grau no qual um dado conteúdo influencia o desenvolvimento futuro de outros conteúdos por todo o cérebro, especialmente no que concerne como esses efeitos se manifestam nos relatos e comportamentos da pessoa em resposta a várias indagações que podem indicar seu estado consciente.

³ A sigla original é “MDM”, de “*Multiple Drafts Model*”.

Uma das afirmações principais do MMC é que indagações diferentes (por exemplo, receber diferentes questões ou estar em contextos diferentes que fazem demandas comportamentais diferentes) podem levar a respostas diferentes acerca do estado consciente da pessoa. Ademais, de acordo com o MMC, pode ser que não haja uma questão de fato, independente da investigação, sobre qual era, efetivamente, o estado consciente da pessoa. Daí o “múltiplo” de Modelos de Múltiplas Camadas.

O MMC é representacionalista, pois analisa a consciência em termos de relações de conteúdo. Ele também nega a existência dos qualia e, portanto, rejeita qualquer tentativa de distinguir estados conscientes de estados não-conscientes por meio de sua presença. Rejeita, também, a noção do eu como um observador interno, esteja ele localizado no Teatro Cartesiano ou em qualquer outro lugar. O MMC trata o eu como um aspecto emergente ou virtual da narrativa aproximadamente serial e coerente que é construída pelo jogo interativo dos conteúdos no sistema. Muitos desses conteúdos estão interligados no nível intencional, como percepções ou fixações de um ponto de vista relativamente unificado e estendido temporalmente, ou seja, são coerentes em seus conteúdos como se fossem as experiências de um eu em movimento. Mas é a ordem de dependência que é crucial para a abordagem de MMC. Os conteúdos relevantes não são unificados porque são todos observados por um único eu, mas exatamente o contrário. É porque são unificados e coerentes no nível do conteúdo que eles contam como experiências de um eu único, ao menos de um eu virtual único.

É nesse aspecto que o MMC compartilha alguns elementos com as teorias de ordem superior. Os conteúdos que compõem a narrativa em série são, ao menos implicitamente, aqueles de um eu em fluxo, mesmo que virtual, e são eles que são mais provavelmente expressos nos relatos que a pessoa faz de seu estado consciente, em resposta às várias indagações. Eles envolvem, então, um certo grau de reflexividade ou autoconsciência de um tipo que é central para as teorias de ordem superior, mas o aspecto de ordem superior é mais uma característica implícita da corrente de conteúdos do que presente em estados explícitos e distintos de ordem superior, do tipo encontrado em teorias OS convencionais.

O MMC de Dennett é altamente influente, mas suscitou também críticas, principalmente por parte daqueles que o acharam insuficientemente realista na sua abordagem da consciência e, na melhor das hipóteses, incompleto quanto à execução de seu proclamado objetivo de explicá-la completamente (Block 1994, Dretske 1994, Levine 1994). Muitos de seus críticos reconhecem o valor e *insight* do MMC, mas negam que não haja outros fatos reais sobre a consciência além daqueles capturados pela teoria (Rosenthal 1994, Van Gulick 1994, Akins 1996).

Um grande exemplo psicológico da abordagem cognitiva é a teoria do ambiente de trabalho global, de Bernard Baars (1988). Ela oferece um modelo psicológico e, em grande medida, funcional da consciência, que lida mais diretamente com a noção de consciência de acesso e tem muito em comum com a ideia da consciência como uma forma de celebridade cerebral, do modelo de múltiplas camadas. A principal ideia das teorias de ambiente de trabalho global é que a consciência é uma capacidade ou módulo de recurso limitado que disponibiliza a informação a ser “transmitida” amplamente por todo o sistema, e que permite um processamento sofisticado mais flexível. Ela é, assim, bastante similar aos vários modelos em psicologia cognitiva sobre a atenção e a memória de trabalho. Muitos outros modelos cognitivos da consciência baseados psicologicamente foram propostos, incluindo aqueles de George Mandler (1975), Timothy Shallice (1988), e Daniel Schacter (1989).

9.4 Teorias neurais

Teorias neurais da consciência assumem várias formas, ainda que de alguma maneira a maioria se refira aos chamados “correlatos neurais da consciência”, ou CNC.⁴ A menos que se seja um dualista ou um outro tipo de não-fisicalista, algo mais do que mera correlação é necessário; ao menos alguns CNC precisam ser os substratos essenciais da consciência. Uma teoria neural informativa precisa explicar por que ou como as correlações relevantes existem, e se a teoria está comprometida com o fisicalismo, precisará mostrar como os substratos

⁴ A sigla original é “NCCs”, de “*Neural Correlates of Consciousness*”.

neurais subjacentes podem ser idênticos a estados conscientes, ou ao menos realizá-los ao executar suas funções ou condições exigidas (Metzinger 2000).

Essas teorias são diversificadas, não apenas quanto aos processos ou propriedades neurais a que apelam, mas também quanto aos aspectos da consciência que tomam como seus respectivos *explananda*. Algumas são baseadas em características sistêmicas de alto nível do cérebro, mas outras se concentram em propriedades fisiológicas ou estruturais mais específicas, com as correspondentes diferenças em suas pretendidas metas explicativas.

Uma amostra de teorias neurais recentes deveria incluir modelos que apelam para campos integrados globais (Kinsbourne), ligações através de oscilações sincrônicas (Singer 1999, Crick e Koch 1990), conjunções neurais transitórias mediadas por NMDA (Flohr 1995), padrões talâmicos modulados de ativação cortical (Llinas 2001), circuitos corticais de re-entrada (Edelman 1989), mecanismos comparativos que executam círculos contínuos de ação-predição-avaliação entre as áreas frontais e o mesencéfalo (Gray 1995), hemisfério esquerdo baseado em processos interpretativos (Gazzaniga 1988), e processos hemostáticos somatossensórios emotivos baseados nonexo frontal-límbico (Damásio 1999), ou na substância cinzenta periaquedutal (Panksepp 1998).

Em cada caso, o objetivo é explicar como a organização e a atividade no nível neural relevante poderiam estar por trás de uma ou outra característica ou tipo importante de consciência. Campos globais ou conjunções sincrônicas transitórias poderiam estar por trás da unidade intencional da consciência fenomenal. Plasticidade baseada em NMDA, projeções talâmicas específicas no córtex ou ondas oscilatórias regulares poderiam todas contribuir para a formação de padrões ou regularidades neurais de curto prazo, mas bem difundidas, necessárias para unir experiências conscientes integradas fora da atividade local, em diversos módulos cerebrais especializados. Processos interpretativos do hemisfério esquerdo poderiam oferecer uma base para formas narrativas de autoconsciência. Assim, é possível que múltiplas teorias neurais distintas sejam todas verdadeiras, cada uma contribuindo com algum entendimento parcial das ligações entre a mentalidade consciente, em suas diversas formas, e o cérebro ativo, em seus muitos níveis de organização e estrutura complexas.

9.5 Teorias quânticas

Outras teorias físicas foram além do neural e situaram o lugar natural da consciência num nível bem mais fundamental: em particular, no nível micro-físico dos fenômenos quânticos. De acordo com essas teorias, a natureza e a base da consciência não podem ser adequadamente compreendidas dentro do quadro da física clássica, mas devem ser buscadas dentro do quadro alternativo da realidade física fornecido pela mecânica quântica. Os proponentes da abordagem da consciência quântica veem a natureza radicalmente alternativa e muitas vezes contra-intuitiva da física quântica como exatamente o que é preciso para superar os supostos obstáculos explicativos que confrontam tentativas mais convencionais de eliminar a lacuna psico-física.

Mais uma vez, há um amplo leque de teorias e modelos específicos que foram propostos apelando para vários fenômenos quânticos, de forma a explicar várias características da consciência. Seria impossível catalogá-los aqui e mesmo explicar, de uma maneira substancial, as principais características da mecânica quântica às quais apelam. Entretanto, um breve exame seletivo pode prover um sentido, ainda que parcial e obscuro, das opções que foram propostas.

O físico Roger Penrose (1989, 1994) e o anesthesiologista Stuart Hameroff (1998) encabeçaram um modelo segundo o qual a consciência surge através de efeitos quânticos que ocorrem dentro de estruturas sub-celulares internas aos neurônios, chamadas de *microtúbulos*. O modelo postula os chamados “*colapsos objetivos*”, que se referem ao sistema quântico mudando de uma superposição de múltiplos estados possíveis para um único estado definido, mas sem a intervenção de um observador ou de uma medida, como na maior parte dos modelos mecânico-quânticos. De acordo com Penrose e Hameroff, o ambiente interno aos microtúbulos é especialmente apropriado para tais colapsos objetivos, e os auto-colapsos resultantes produzem um fluxo coerente que regula a atividade neuronal e torna possível processos mentais não-algorítmicos.

O psiquiatra Ian Marshall propôs um modelo que pretende explicar a unidade coerente da consciência apelando para a produção, no cérebro, de um estado físico parecido com um *condensado de Bose-Einstein*. Este é um fenômeno quântico no qual uma coleção de átomos age como uma única entidade coerente, e a distinção entre átomos discretos se perde. Apesar de estados cerebrais não serem, literalmente, exemplos de condensados de Bose-Einstein, razões foram dadas para mostrar por que cérebros são mais suscetíveis a gerar estados capazes de exibir uma coerência similar (Marshall e Zohar 1990).

Uma base para a consciência também foi buscada na natureza *holista* da mecânica quântica e no fenômeno do *entrelaçamento*, de acordo com o qual partículas que interagiram continuam tendo suas naturezas dependentes uma da outra, mesmo depois da sua separação. Não surpreendentemente, esses modelos pretendem explicar principalmente a coerência da consciência, mas também foram invocados como um desafio mais geral à concepção atomista da física tradicional, segundo a qual as propriedades do todo são explicáveis por apelo às propriedades das partes mais seu modo de combinação, um método de explicação que hoje poderia ser visto como fracassado para explicar a consciência (Silberstein 1998, 2001).

Outros viram a mecânica quântica como indicando que a consciência é uma propriedade absolutamente fundamental da realidade física, uma propriedade que precisa ser introduzida no nível mais básico de todos (Stapp 1993). Apelaram principalmente para o papel do observador no colapso da função de onda, isto é, o colapso da realidade quântica de uma superposição de estados possíveis em um único estado definido, quando a medição é feita. Esses modelos podem ou não adotar uma forma de quase-idealismo, na qual a própria existência da realidade física depende dela ser conscientemente observada.

Existem muitos outros modelos quânticos da consciência que podem ser encontrados na literatura – alguns defendendo uma metafísica radicalmente revisionista, outros não –, mas esses quatro fornecem uma amostra razoável, ainda que parcial, das alternativas.

9.6 Teorias não-físicas

A maior parte das teorias específicas da consciência – sejam cognitivas, neurais ou da mecânica quântica – pretendem explicar ou propor modelos da consciência vista como uma característica natural do mundo físico. No entanto, aqueles que rejeitam uma ontologia fisicalista da consciência precisam encontrar formas alternativas de tratá-la como um aspecto não-físico da realidade. Portanto, aqueles que adotam uma visão metafísica dualista ou antifisicalista precisam, por fim, fornecer modelos específicos da consciência, diferentes dos cinco tipos mencionados. Tanto os dualistas de substância quanto os dualistas de propriedades precisam desenvolver os detalhes de suas teorias de forma a articular as naturezas específicas das características não-físicas relevantes da natureza, com as quais igualam a consciência ou às quais apelam para explicá-la.

Uma grande variedade desses modelos foi proposta, incluindo os seguintes. David Chalmers (1996) ofereceu uma versão admitidamente especulativa de pansiquismo, que apela para a noção de informação, não apenas para explicar invariâncias psicofísicas entre o fenomenal e espaços informacionais fisicamente realizados, mas possivelmente também para explicar a ontologia do físico como em si mesma derivada do informacional (uma versão da teoria do *“it from bit”*). De forma um tanto similar, Gregg Rosenberg propôs recentemente (2004) uma abordagem da consciência que, simultaneamente, trata a base categorial última das relações causais. Tanto no caso causal quanto no caso consciente, Rosenberg argumenta que fatos funcionais e relacionais precisam, no fim das contas, depender de uma base categorial não-relacional, e oferece um modelo segundo o qual relações causais e fatos fenomenais qualitativos ambos dependem da mesma base. Ademais, como mencionado anteriormente (seção 9.5), algumas teorias quânticas tratam a consciência como uma característica fundamental da realidade (Stapp 1993) e, na medida em que assim o fazem, também poderiam ser plausivelmente classificadas como teorias não-físicas.

10. Conclusão

Uma compreensão completa da consciência provavelmente exigirá teorias de vários tipos. Pode-se, com utilidade e sem contradição, aceitar uma diversidade de modelos que, cada um à sua maneira, buscam explicar, respectivamente, os aspectos físicos, neurais, cognitivos, funcionais, representacionais e de ordem superior da consciência. É improvável que haja uma perspectiva teórica única que seja suficiente para explicar todas as características da consciência que pretendemos compreender. Portanto, uma abordagem sintética e pluralista deve oferecer a melhor via para futuros progressos.

Bibliografia

- Akins, K. 1993. "A bat without qualities?" In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Akins, K. 1996. "Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness." *Mind and Language*, 11: 1-43.
- Anderson, J. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Armstrong, D. 1968. *A Materialist Theory of Mind*, London: Routledge and Kegan Paul.
- Armstrong, D. 1981. "What is consciousness?" In *The Nature of Mind*. Ithaca, NY: Cornell University Press.
- Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Balog, K. 1999. "Conceivability, possibility, and the mind-body problem." *Philosophical Review*, 108: 497-528.
- Block, N. 1980a. "Troubles with Functionalism," in *Readings in the Philosophy of Psychology*, Volume 1, Ned Block, ed., Cambridge, MA : Harvard University Press, 268-305.
- Block, N. 1980b. "Are absent qualia impossible?" *Philosophical Review*, 89/2: 257-74.

- Block, N. 1990. "Inverted Earth," *Philosophical Perspectives*, 4, J. Tomberlin, ed., Atascadero, CA: Ridgeview Publishing Company.
- Block, N. 1995. "On a confusion about the function of consciousness." *Behavioral and Brain Sciences*, 18: 227-47.
- Block, N. 1994. "What is Dennett's theory a theory of?" *Philosophical Topics*, 22/1-2: 23-40.
- Block, N. 1996. "Mental paint and mental latex." In E. Villanueva, ed. *Perception*. Atascadero, CA: Ridgeview.
- Block, N. and Stalnaker, R. 1999. "Conceptual analysis, dualism, and the explanatory gap." *Philosophical Review*, 108/1: 1-46.
- Boyd, R. 1980. "Materialism without reductionism: What physicalism does not entail." In N. Block, ed. *Readings in the Philosophy of Psychology*, Vol. 1. Cambridge, MA: Harvard University Press.
- Brook, A. and Raymont, P. 2010. "The Unity of Consciousness", *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/consciousness-unity/>>.
- Byrne, A. 1997. "Some like it HOT: consciousness and higher-order thoughts." *Philosophical Studies*, 2: 103-29.
- Byrne, A. 2001. "Intentionalism defended". *Philosophical Review*, 110: 199-240.
- Campbell, K. 1970. *Body and Mind*. New York: Doubleday.
- Campbell, J. 1994. *Past, Space, and Self*. Cambridge, MA: MIT Press.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Carruthers, P. 2011. "Higher-Order Theories of Consciousness", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/consciousness-higher/>>.

- Chalmers, D. 1995. "Facing up to the problem of consciousness". *Journal of Consciousness Studies*, 2: 200-19.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. 2002. "Does conceivability entail possibility?" In T. Gendler and J. Hawthorne eds. *Conceivability and Possibility*. Oxford: Oxford University Press.
- Chalmers, D. 2003. "The content and epistemology of phenomenal belief." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Chalmers, D. and Jackson, F. 2001. "Conceptual analysis and reductive explanation". *Philosophical Review*, 110/3: 315-60.
- Churchland, P. M. 1985. "Reduction, qualia, and direct introspection of brain states". *Journal of Philosophy*, 82: 8-28.
- Churchland, P. M. 1995. *The Engine of Reason and Seat of the Soul*. Cambridge, MA: MIT Press.
- Churchland, P. S. 1981. "On the alleged backwards referral of experiences and its relevance to the mind body problem". *Philosophy of Science*, 48: 165-81.
- Churchland, P. S. 1983. "Consciousness: the transmutation of a concept". *Pacific Philosophical Quarterly*, 64: 80-95.
- Churchland, P. S. 1996. "The hornswoggle problem". *Journal of Consciousness Studies*, 3: 402-8.
- Clark, A. 1993. *Sensory Qualities*. Oxford: Oxford University Press.
- Clark, G. and Riel-Salvatore, J. 2001. "Grave markers, middle and early upper paleolithic burials". *Current Anthropology*, 42/4: 481-90.
- Cleeremans, A., ed. 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press.
- Crick, F. and Koch, C. 1990. "Toward a neurobiological theory of consciousness". *Seminars in Neuroscience*, 2: 263-75.

- Crick, F. H. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribners.
- Davies, M. and Humphreys, G. 1993. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Damasio, A. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt.
- Dennett, D. C. 1978. *Brainstorms*. Cambridge: MIT Press.
- Dennett, D. C. 1984. *Elbow Room: The Varieties of Free Will Worth Having*. Cambridge: MIT Press.
- Dennett, D. C. 1990. "Quining qualia". In *Mind and Cognition*, W. Lycan, ed., Oxford : Blackwell, 519-548.
- Dennett, D. C. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Dennett, D. C. 1992. "The self as the center of narrative gravity". In F. Kessel, P. Cole, and D. L. Johnson, eds. *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Lawrence Erlbaum.
- Dennett, D. C. 2003. *Freedom Evolves*. New York: Viking.
- Dennett, D. C. and Kinsbourne, M. 1992. "Time and the observer: the where and when of consciousness in the brain". *Behavioral and Brain Sciences*, 15: 187-247.
- Descartes, R. 1644/1911. *The Principles of Philosophy*. Translated by E. Haldane and G. Ross. Cambridge: Cambridge University Press.
- Dretske, F. 1993. "Conscious experience." *Mind*, 102: 263-283.
- Dretske, F. 1994. "Differences that make no difference". *Philosophical Topics*, 22/1-2: 41-58.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, Mass: The MIT Press, Bradford Books.
- Eccles, J. and Popper, K. 1977. *The Self and Its Brain: An Argument for Interactionism*. Berlin: Springer

- Edelman, G. 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Farah, M. 1990. *Visual Agnosia*. Cambridge: MIT Press.
- Flanagan, O. 1992. *Consciousness Reconsidered*. Cambridge, MA: MIT Press.
- Flohr, H. 1995. "An information processing theory of anesthesia". *Neuropsychologia*, 33/9: 1169-80.
- Flohr, H., Glade, U. and Motzko, D. 1998. "The role of the NMDA synapse in general anesthesia". *Toxicology Letters*, 100-101: 23-29.
- Fodor, J. 1974. "Special sciences". *Synthese*, 28: 77-115.
- Fodor, J. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Foster, J. 1989. "A defense of dualism". In J. Smythies and J. Beloff, eds. *The Case for Dualism*. Charlottesville, VA: University of Virginia Press.
- Foster J. 1996. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of Mind*. London: Routledge.
- Gallistel, C. 1990. *The Organization of Learning*. Cambridge, MA: MIT Press.
- Gardiner, H. 1985. *The Mind's New Science*. New York: Basic Books.
- Gazzaniga, M. 1988. *Mind Matters: How Mind and Brain Interact to Create our Conscious Lives*. Boston: Houghton Mifflin.
- Gennaro, R. 1995. *Consciousness and Self-consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- Gennaro, R., ed. 2004. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.

- Gertler, B. 2011. "Self-Knowledge", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/self-knowledge/>.
- Gray, J. 1995. "The contents of consciousness: a neuropsychological conjecture". *Behavior and Brain Sciences*, 18/4: 659-722.
- Hameroff, S. 1998. "Quantum computation in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness". *Philosophical Transactions Royal Society London*, A 356: 1869-96.
- Hardin, C. 1986. *Color for Philosophers*. Indianapolis: Hackett.
- Hardin, C. 1992. "Physiology, phenomenology, and Spinoza's true colors". In A. Beckermann, H. Flohr, and J. Kim, eds. *Emergence or Reduction?: Prospects for Nonreductive Physicalism*. Berlin and New York: De Gruyter.
- Harman, G. 1990. "The intrinsic quality of experience". In J. Tomberlin, ed. *Philosophical Perspectives*, 4. Atascadero, CA: Ridgeview Publishing.
- Hartshorne, C. 1978. "Panpsychism: mind as sole reality". *Ultimate Reality and Meaning*, 1: 115-29.
- Hasker, W. 1999. *The Emergent Self*. Ithaca, NY: Cornell University Press.
- Heidegger, M. 1927/1962. *Being and Time (Sein und Zeit)*. Translated by J. Macquarrie and E. Robinson. New York: Harper and Row.
- Hellman, G. and Thompson, F. 1975. "Physicalism: ontology, determination and reduction". *Journal of Philosophy*, 72: 551-64.
- Hill, C. 1991. *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.
- Hill, C. 1997. "Imaginability, conceivability, possibility, and the mind-body problem". *Philosophical Studies*, 87: 61-85.

- Hill, C. and McLaughlin, B. 1998. "There are fewer things in reality than are dreamt of in Chalmers' philosophy". *Philosophy and Phenomenological Research*, 59/2: 445-54.
- Horgan, T. 1984. "Jackson on Physical Information and Qualia." *Philosophical Quarterly*, 34: 147-83.
- Horgan, T. and Tienson, J. 2002. "The intentionality of phenomenology and the phenomenology of intentionality". In D. J. Chalmers , ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.
- Hume, D. 1739/1888. *A Treatise of Human Nature*. ed. L Selby-Bigge. Oxford: Oxford University Press.
- Humphreys, N. 1982. *Consciousness Regained*. Oxford: Oxford University Press.
- Humphreys, N. 1992. *A History of the Mind*. London: Chatto and Windus.
- Husserl, E. 1913/1931. *Ideas: General Introduction to Pure Phenomenology (Ideen au einer reinen Phänomenologie und phänomenologischen Philosophie)*. Translated by W. Boyce Gibson. New York: MacMillan.
- Husserl, E. 1929/1960. *Cartesian Meditations: an Introduction to Phenomenology*. Translated by Dorian Cairns. The Hague: M. Nijhoff.
- Huxley, T. 1866. *Lessons on Elementary Physiology* 8. London
- Huxley, T. 1874. "On the hypothesis that animals are automata". *Fortnightly Review*, 95: 555-80. Reprinted in *Collected Essays*. London, 1893.
- Hurley, S. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Jackson, F. 1982. "Epiphenomenal qualia". *Philosophical Quarterly*, 32: 127-136.
- Jackson, F. 1986. "What Mary didn't know". *Journal of Philosophy*, 83: 291-5.
- Jackson, F. 1993. "Armchair metaphysics". In J. O'Leary-Hawthorne and M. Michael, eds. *Philosophy of Mind*. Dordrecht: Kluwer Books.

- Jackson, F. 1998. "Postscript on qualia". In F. Jackson *Mind, Method and Conditionals*. London: Routledge.
- Jackson, F. 2004. "Mind and illusion." In P. Ludlow, Y. Nagasawa and D. Stoljar eds. *There's Something about Mary: Essays on the Knowledge Argument*. Cambridge, MA: MIT Press.
- James, W. 1890. *The Principles of Psychology*. New York: Henry Holt and Company.
- Jaynes, J. 1974. *The Origins of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin.
- Kim, J. 1987. "The myth of non-reductive physicalism". *Proceedings and Addresses of the American Philosophical Association*.
- Kim, J. 1998. *Mind in Physical World*. Cambridge: MIT Press.
- Kirk, R. 2012. "Zombies", *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/spr2012/entries/zombies/>.
- Kant, I. 1787/1929. *Critique of Pure Reason*. Translated by N. Kemp Smith. New York: MacMillan.
- Kinsbourne, M. 1988. "Integrated field theory of consciousness". In A. Marcel and E. Bisiach, eds. *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Kirk, R. 1974. "Zombies vs materialists". *Proceedings of the Aristotelian Society*, Supplementary Volume, 48: 135-52.
- Kirk, R. 1991. "Why shouldn't we be able to solve the mind-body problem?" *Analysis*, 51: 17-23.
- Klima, G. 2008. "The Medieval Problem of Universals", *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2008/entries/universals-medieval/>.
- Köhler, W. 1929. *Gestalt Psychology*. New York: Liveright.

- Köffka, K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt Brace.
- Leibniz, G. W. 1686 /1991. *Discourse on Metaphysics*. Translated by D. Garter and R. Aries. Indianapolis: Hackett.
- Leibniz, G. W. 1720/1925. *The Monadology*. Translated by R. Lotte. London: Oxford University Press.
- Levin, J. 2010. "Functionalism", *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2010/entries/functionalism/>.
- Levine, J. 1983. "Materialism and qualia: the explanatory gap". *Pacific Philosophical Quarterly*, 64: 354-361.
- Levine, J. 1993. "On leaving out what it's like". In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Levine, J. 1994. "Out of the closet: a qualophile confronts qualophobia". *Philosophical Topics*, 22/1-2: 107-26.
- Levine, J. 2001. *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, Mass: The MIT Press.
- Lewis, D. 1972. "Psychophysical and theoretical identifications". *Australasian Journal of Philosophy*, 50: 249-58.
- Lewis, D. 1990. "What experience teaches." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell.
- Libet, B. 1982. "Subjective antedating of a sensory experience and mind-brain theories". *Journal of Theoretical Biology*, 114: 563-70.
- Libet, B. 1985. "Unconscious cerebral initiative and the role of conscious will in voluntary action". *Behavioral and Brain Sciences*, 8: 529-66.
- Llinas, R. 2001. *I of the vortex: from neurons to self*. Cambridge, MA: MIT Press

- Loar, B. 1990. "Phenomenal states," in *Philosophical Perspectives*, 4: 81-108.
- Loar, B. 1997. "Phenomenal states". In N. Block, O. Flanagan, and G. Guzeldere eds. *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Locke, J. 1688/1959. *An Essay on Human Understanding*. New York: Dover.
- Lockwood, M. 1989. *Mind, Brain, and the Quantum*. Oxford: Oxford University Press.
- Lorenz, K. 1977. *Behind the Mirror (Rückseite dyes Speigels)*. Translated by R. Taylor. New York: Harcourt Brace Jovanovich.
- Lycan, W. 1987. *Consciousness*. Cambridge, MA: MIT Press.
- Lycan, W. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Lycan, W. 2004. "The superiority of HOP to HOT". In R. Gennaro ed. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- Lycan, W. 2008. "Representational Theories of Consciousness", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/consciousness-representational/>.
- Mandler, G. 1975. *Mind and Emotion*. New York: Wiley.
- Marshall, I. and Zohar, D. 1990. *The Quantum Self: Human Nature and Consciousness Defined by the New Physics*. New York: Morrow.
- McGinn, C. 1989. "Can we solve the mind-body problem?" *Mind*, 98: 349-66
- McGinn, C. 1991. *The Problem of Consciousness*. Oxford: Blackwell.
- McGinn, C. 1995. "Consciousness and space." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Merleau-Ponty, M. 1945/1962. *Phenomenology of Perception (Phénoménologie de lye Perception)*. Translated by Colin Smith. London: Routledge and Kegan Paul.
- Metzinger, T., ed. 1995. *Conscious Experience*. Paderborn: Ferdinand Schöningh.

- Metzinger, T. ed. 2000. *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Mill, J. 1829. *Analysis of the Phenomena of the Human Mind*. London.
- Mill, J.S. 1865. *An Analysis of Sir William Hamilton's Philosophy*. London.
- Moore, G. E. 1922. "The refutation of idealism." In G. E. Moore *Philosophical Studies*. London : Routledge and Kegan Paul.
- Nagel, T. 1974. "What is it like to be a bat?" *Philosophical Review*, 83: 435-456.
- Nagel, T. 1979. "Panpsychism." In T. Nagel *Mortal Questions*. Cambridge: Cambridge University Press.
- Natsoulas, T. 1983. "Concepts of consciousness." *Journal of Mind and Behavior*, 4: 195-232.
- Nelkin, N. 1989. "Unconscious sensations." *Philosophical Psychology*, 2: 129-41.
- Nemirow, L. 1990. "Physicalism and the cognitive role of acquaintance." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell.
- Neisser, U. 1965. *Cognitive Psychology*. Englewood Cliffs: Prentice Hall.
- Nida-Rümelin, M. 1995. "What Mary couldn't know: belief about phenomenal states." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Nida-Rümelin, M. 2010. "Qualia: The Knowledge Argument", *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2010/entries/qualia-knowledge/>>.
- O'Connor, T. 2011. "Free Will", *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2011/entries/freewill/>>.
- Panksepp, J. 1998. *Affective Neuroscience*. Oxford: Oxford University Press.
- Papineau, D. 1994. *Philosophical Naturalism*. Oxford: Blackwell.

- Papineau, D. 1995. "The antipathetic fallacy and the boundaries of consciousness." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Papineau, D. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.
- Peacocke, C. 1983. *Sense and Content*, Oxford: Oxford University Press.
- Pearson, M.P. 1999. *The Archeology of Death and Burial*. College Station, Texas: Texas A&M Press.
- Penfield, W. 1975. *The Mystery of the Mind: a Critical Study of Consciousness and the Human Brain*. Princeton, NJ: Princeton University Press.
- Perry, J. 2001. *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Penrose, R. 1989. *The Emperor's New Mind: Computers, Minds and the Laws of Physics*. Oxford: Oxford University Press.
- Penrose, R. 1994. *Shadows of the Mind*. Oxford: Oxford University Press.
- Place, U. T. 1956. "Is consciousness a brain process?" *British Journal of Psychology*, 44-50.
- Putnam, H. 1975. "Philosophy and our mental life." In H. Putnam *Mind Language and Reality: Philosophical Papers Vol. 2*. Cambridge: Cambridge University Press.
- Putnam, H. and Oppenheim, P. 1958. "Unity of science as a working hypothesis." In H. Fiegl, G. Maxwell, and M. Scriven eds. *Minnesota Studies in the Philosophy of Science II*. Minneapolis: University of Minnesota Press.
- Ramsey, W. 2011. "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/materialism-eliminative/>.
- Rey, G. 1986. "A question about consciousness." In H. Otto and J. Tuedio, eds. *Perspectives on Mind*. Dordrecht: Kluwer.

- Robinson, H. 1982. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge: Cambridge University Press.
- Robinson, H. 2011. "Dualism", *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/dualism/>.
- Robinson, D. 1993. "Epiphenomenalism, laws, and properties." *Philosophical Studies*, 69: 1-34.
- Robinson, W. 2011. "Epiphenomenalism", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/epiphenomenalism/>.
- Rosenberg, G. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.
- Rosenthal, D. 1986. "Two concepts of consciousness." *Philosophical Studies*, 49: 329-59.
- Rosenthal, D. 1991. "The independence of consciousness and sensory quality." In E. Villanueva, ed. *Consciousness*. Atascadero, CA: Ridgeview Publishing.
- Rosenthal, D. M. 1993. "Thinking that one thinks." In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Rosenthal, D. 1994. "First person operationalism and mental taxonomy." *Philosophical Topics*, 22/1-2: 319-50.
- Rosenthal, D. M. 1997. "A theory of consciousness." In N. Block, O. Flanagan, and G. Guzeldere, eds. *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Russell, B. 1927. *The Analysis of Matter*. London: Kegan Paul.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson and Company.
- Sacks, O. 1985. *The Man who Mistook his Wife for a Hat*. New York: Summit.

- Seager, W. 1995. "Consciousness, information, and panpsychism." *Journal of Consciousness Studies*, 2: 272-88.
- Seager, W. e Allen-Hermanson, S. 2012. "Panpsychism", *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/spr2012/entries/panpsychism/>>.
- Schacter, D. 1989. "On the relation between memory and consciousness: dissociable interactions and consciousness." In H. Roediger and F. Craik eds. *Varieties of Memory and Consciousness*. Hillsdale, NJ: Erlbaum.
- Schneider W. and Shiffrin, R. 1977. "Controlled and automatic processing: detection, search and attention." *Psychological Review*, 84: 1-64.
- Searle, J. R. 1990. "Consciousness, explanatory inversion and cognitive science." *Behavioral and Brain Sciences*, 13: 585-642.
- Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
- Siewert, C. 2011. "Consciousness and Intentionality", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/consciousness-intentionality/>>.
- Shallice, T. 1988. *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Shear, J. 1997. *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press.
- Shoemaker, S. 1975. "Functionalism and qualia," *Philosophical Studies*, 27: 291-315.
- Shoemaker, S. 1981. "Absent qualia are impossible." *Philosophical Review*, 90: 581-99.
- Shoemaker, S. 1982. "The inverted spectrum." *Journal of Philosophy*, 79: 357-381.

- Shoemaker, S. 1990. "Qualities and qualia: what's in the mind," *Philosophy and Phenomenological Research*, Supplement, 50: 109-131.
- Shoemaker, S. 1998. "Two cheers for representationalism," *Philosophy and Phenomenological Research*.
- Silberstein, M. 1998. "Emergence and the mind-body problem." *Journal of Consciousness Studies*, 5: 464-82.
- Silberstein, M. 2001. "Converging on emergence: consciousness, causation and explanation." *Journal of Consciousness Studies*, 8: 61-98.
- Singer, P. 1975. *Animal Liberation*. New York: Avon Books.
- Singer, W. 1999. "Neuronal synchrony: a versatile code for the definition of relations." *Neuron*, 24: 49-65.
- Skinner, B. F. 1953. *Science and Human Behavior*. New York: MacMillan.
- Smart, J. 1959. "Sensations and brain processes." *Philosophical Review*, 68: 141-56.
- Stapp, H. 1993. *Mind, Matter and Quantum Mechanics*. Berlin: Springer Verlag.
- Stoljar, D. 2001. "Two conceptions of the physical." *Philosophy and Phenomenological Research*, 62: 253-81
- Strawson, G. 1994. *Mental Reality*. Cambridge, Mass: MIT Press, Bradford Books.
- Swinburne, R. 1986. *The Evolution of the Soul*. Oxford: Oxford University Press.
- Titchener, E. 1901. *An Outline of Psychology*. New York: Macmillan.
- Travis, C. 2004. "The silence of the senses." *Mind*, 113: 57-94.
- Triesman, A. and Gelade, G. 1980. "A feature integration theory of attention." *Cognitive Psychology*, 12: 97-136.
- Tye, M. 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.

Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Tye, M. 2003. "Blurry images, double vision and other oddities: new troubles for representationalism?" In A. Jokic and Q. Smith eds., *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.

Van Gulick, R. 1985. "Physicalism and the subjectivity of the mental." *Philosophical Topics*, 13: 51-70.

Van Gulick, R. 1992. "Nonreductive materialism and intertheoretical constraint." In A. Beckermann, H. Flohr, J. Kim, eds. *Emergence and Reduction*. Berlin and New York: De Gruyter, 157-179.

Van Gulick, R. 1993. "Understanding the phenomenal mind: Are we all just armadillos?" In M. Davies and G. Humphreys, eds., *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.

Van Gulick, R. 1994. "Dennett, drafts and phenomenal realism." *Philosophical Topics*, 22/1-2: 443-56.

Van Gulick, R. 1995. "What would count as explaining consciousness?" In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.

Van Gulick, R. 2000. "Inward and upward: reflection, introspection and self-awareness." *Philosophical Topics*, 28: 275-305.

Van Gulick, R. 2003. "Maps, gaps and traps." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.

Van Gulick, R. 2004. "Higher-order global states HOGS: an alternative higher-order model of consciousness." In Gennaro, R. ed. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.

van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.

Varela, F. and Maturana, H. 1980. *Cognition and Autopoiesis*. Dordrecht: D. Reidel.

- Varela, F. 1995. "Neurophenomenology: A methodological remedy for the hard problem." *Journal of Consciousness Studies*, 3: 330-49.
- Varela, F. and Thomson, E. 2003. "Neural synchronicity and the unity of mind: a neurophenomenological perspective." In Cleermans, A. ed. *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford: Oxford University Press
- Velmans, M. 1991. "Is Human information processing conscious?" *Behavioral and Brain Sciences*, 14/4: 651-668
- Velmans, M. 2003. "How could conscious experiences affect brains?" *Journal of Consciousness Studies*, 9: 3-29.
- von Helmholtz, H. 1897/1924. *Treatise on Physiological Optics*. Translated by J. Soothly. New York: Optical Society of America.
- Wilkes, K. V. 1984. "Is consciousness important?" *British Journal for the Philosophy of Science*, 35: 223-43.
- Wilkes, K. V. 1988. "Yishi, duo, us and consciousness." In A. Marcel and E. Bisiach, eds., *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Wilkes, K. V. 1995. "Losing consciousness." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Watson, J. 1924. *Behaviorism*. New York: W. W. Norton.
- Wegner, D. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wittgenstein, L. 1921/1961. *Tractatus Logico-Philosophicus*. Translated by D. Pears and B. McGuinness. London: Routledge and Kegan Paul.
- Wundt, W. 1897. *Outlines of Psychology*. Leipzig: W. Engleman.
- Yablo, S. 1998. "Concepts and consciousness." *Philosophy and Phenomenological Research*, 59: 455-63.